



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - KS 141501

**PREDIKSI INTENSITAS HUJAN KOTA
SURABAYA DENGAN MATLAB
MENGUNAKAN TEKNIK RANDOM FOREST
DAN CART (STUDI KASUS KOTA SURABAYA)**

MAULANA DHAWANGKHARA
NRP 5213 100 074

Dosen Pembimbing
Edwin Riksakomara, S.Kom, MT

JURUSAN SISTEM INFORMASI
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember
Surabaya 2016



FINAL PROJECT - KS 141501

***RAINFALL INTENSITY PREDICTION WITH
MATLAB USING CART AND RANDOM
FOREST METHOD (CASE STUDY:
SURABAYA)***

MAULANA DHAWANGKHARA
NRP 5213 100 074

Supervisor
Edwin Riksakomara, S.Kom, MT

INFORMATION SYSTEMS DEPARTEMENT
Faculty of Information Technology
Institut Teknologi Sepuluh Nopember
Surabaya 2016

LEMBAR PENGESAHAN

PREDIKSI INTENSITAS HUJAN KOTA SURABAYA DENGAN MATLAB MENGGUNAKAN TEKNIK RANDOM FOREST DAN CART (STUDI KASUS KOTA SURABAYA)

TUGAS AKHIR

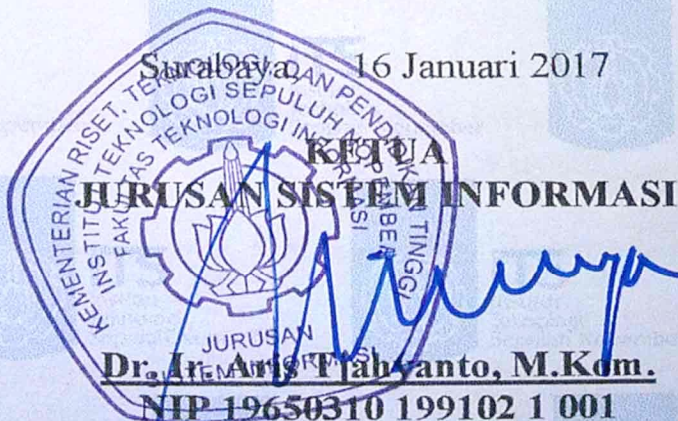
Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada

Jurusan Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh:

MAULANA DHAWANGKHARA
5213 100 074

Surabaya, 16 Januari 2017



LEMBAR PERSETUJUAN

PREDIKSI INTENSITAS HUJAN KOTA SURABAYA DENGAN MATLAB MENGGUNAKAN TEKNIK RANDOM FOREST DAN CART (STUDI KASUS KOTA SURABAYA)

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

pada

Jurusan Sistem Informasi

Fakultas Teknologi Informasi

Institut Teknologi Sepuluh Nopember

Oleh :

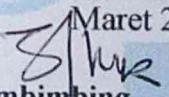
MAULANA DHAWANGKHARA

5213 100 074

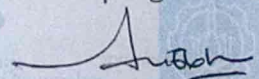
Disetujui Tim Penguji : Tanggal Ujian : 10 Januari 2017

Periode Wisuda: Maret 2017

Edwin Riksakomara, S.Kom, MT


(Pembimbing 1)

Wiwik Anggraeni S.Si., M.Kom


(Penguji 1)

Faisal Mahananto, S.Kom, M.Eng, Ph.D


(Penguji 2)

PREDIKSI INTENSITAS HUJAN KOTA SURABAYA DENGAN MATLAB MENGGUNAKAN TEKNIK RANDOM FOREST DAN CART (STUDI KASUS KOTA SURABAYA)

Nama Mahasiswa : Maulana Dhawangkharu
NRP : 5213 100 074
Jurusan : SISTEM INFORMASI FTIF-ITS
Dosen Pembimbing : Edwin Riksakomara, S.Kom, MT

ABSTRAK

Keakuratan prediksi potensi curah hujan di Kota Surabaya dibutuhkan untukantisipasi bencana akibat hujan seperti banjir bandang, membantu memprediksi kondisi penerbangan dan membantu majaemen saluran sanitasi di Surabaya. Prediksi dilakukan dengan data hari sebelumnya menggunakan perbandingan teknik Classification and Regression Trees (CART) dan Random Forest (RF) pada data cuaca selama 17 tahun (2000-2016) berasal dari stasiun cuaca Juanda, Surabaya melalui website NCDC (National Climate Data Center) yang terdiri dari data suhu udara, titik embun, kecepatan angin, tekanan udara, visibilitas dan curah hujan.

Pencarian model dengan kedua metode di atas dilakukan dengan mencari model yang paling optimal untuk masing-masing metode dengan melakukan pemilihan variabel yang berpengaruh, penyetelan (tuning) parameter metode lalu pengujian dengan proporsi data training dan testing. Perbandingan hasil performa kedua model juga ditinjau untuk mendapatkan model prediktif terbaik. Metode CART dan Random Forest tersebut kemudian diimplementasikan dengan perangkat lunak yang dibangun dengan matlab.

Penggunaan parameter default sesuai Breimann menunjukkan metode Random forest memiliki nilai performa sedikit lebih baik dibandingkan dengan CART sebesar 6%. Eksperimen

tuning parameter untuk kedua metode membuktikan performa yang lebih baik dibandingkan parameter default metode, dengan akurasi baik sebesar 78% untuk 4 dari 5 kelas intensitas hujan, dengan kelas terakhir belum mampu diklasifikasi oleh kedua metode dan mampu memberikan kestabilan hasil performa dari segi uji coba proporsi data training dan testing. Variabel yang berpengaruh besar dalam model CART dan random forest dengan nilai uji performa yang baik antara lain adalah suhu udara, titik embun, suhu udara maksimum dan suhu udara minimum beserta variabel turunannya (selisih suhu udara maksimum dan minimum, selisih suhu udara dan titik embun dan kelembapan relatif).

Tugas akhir ini menghasilkan aplikasi prediksi intensitas hujan yang memiliki kemampuan prediktif yang mampu memberikan akurasi tinggi atas hasil prediksi intensitas hujan (tidak hujan, ringan, sedang, deras, sangat deras).

Kata Kunci: CART, Random Forest, RF, Matlab, Klasifikasi, Curah Hujan, Intensitas Hujan

RAINFALL INTENSITY PREDICTION WITH MATLAB USING CART AND RANDOM FOREST METHOD (CASE STUDY: SURABAYA)

Name : Maulana Dhawangkhara
NRP : 5213 100 074
Department : INFORMATION SYSTEMS FTIF-ITS
Supervisor : Edwin Riksakomara, S.Kom, MT

ABSTRACT

Accuracy of the rainfall prediction in Surabaya most likely needed for anticipating disasters due to rain such as flash floods, predicting the flight conditions and managing sanitation channel in Surabaya. Prediction is done using data of the previous day using Classification and Regression Trees (CART) and Random Forest (RF) method and 17 year weather data (2000-2016) collected from weather stations Juanda, Surabaya through NCDC (National Climatic Data Center). The data consists of air temperature, dew point, wind speed, air pressure, visibility and precipitation.

Model with both the above method is done by searching the most optimal model for each method by adjusting (tuning) the parameters. Comparison of the performance results of both models are also analyzed to get the best classification model.

Construction of classification model with accuracy, precision and recall as measurements showed that both CART and Random Forest method are capable of classifying 4 classes (0,1,2,3) out of 5 classes (0,1,2,3,4) with accuracy of 78%, with the class 4 has not been able to be classified by both methods. Random forest model has slightly better performance than CART at 6%.

Experiments by tuning parameters for both methods prove a better performance than the default parameters and capable

of providing stability performance results in terms of the proportion of data training and data testing. Variables that have great influence in the CART and random forest model are air temperature, dew point, maximum air temperature and minimum air temperature along with 4 variable derivatives (the difference between air temperature maximum and minimum, the difference between the air temperature and the dew point and relative humidity).

This final project sought to develop a Matlab-based application that generates predictions of rainfall intensity that is able to provide high accuracy on the rain intensity (no rain, light, medium, heavy, very heavy).

Keywords: CART, Random Forest, RF, Matlab, Classification, Rainfall, Rainfall Intensity

KATA PENGANTAR

Puji syukur penulis panjatkan atas kehadiran Tuhan Yang Maha Esa atas segala berkat dan rahmat-Nya lah penulis dapat menyelesaikan buku tugas akhir dengan judul “**PREDIKSI INTENSITAS HUJAN KOTA SURABAYA DENGAN MATLAB MENGGUNAKAN TEKNIK RANDOM FOREST DAN CART (STUDI KASUS KOTA SURABAYA)**” yang merupakan salah satu syarat kelulusan pada Jurusan Sistem Informasi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember Surabaya.

Secara khusus penulis akan menyampaikan ucapan terima kasih yang sedalam-dalamnya kepada:

1. Ibu Nur Hasanah dan Bapak Hari Tariyadi selaku kedua orang tua, serta segenap keluarga penulis yang selalu memberikan dukungan dan motivasi. Terima kasih atas doa dan dukungannya yang terus mengalir tiada henti.
2. Bapak Edwin Riksakomara, S.Kom, MT selaku dosen pembimbing dengan penuh keikhlasan dan dedikasi tinggi telah membimbing penulis dalam mengerjakan tugas akhir ini hingga selesai. Terima kasih atas kesediaan, waktu, dan ilmu yang telah diberikan.
3. Ibu Wiwik Anggraeni S.Si., M.Kom dan Bapak Faisal Mahananto, S.Kom, M.Eng, Ph.D selaku dosen penguji yang selalu memberikan saran dan masukan guna kebaikan Tugas Akhir ini.
4. Para teman dekat yang selalu menemani dan memberikan dukungan pada penulis selama masa kuliah: Pejuang FRS (Niko, Valliant, Tetha, Rica, Agung, Juki, Kusnanta, Asvin, Egan).
5. Untuk seluruh teman-teman SI 2013 (Beltranis) yang kerap menjadi tempat berbagi keluh kesah dan dukungan selama kuliah.

6. Seluruh dosen pengajar, staff, dan karyawan di Jurusan Sistem Informasi FTIF ITS Surabaya yang telah memberikan ilmu dan bantuan kepada penulis selama ini.
7. Manga artist yang menginspirasi penulis, Oda-sensei dan Kubo-sensei. Tokoh yang menginspirasi: Akatsuki, Genei Ryodan, CP9, Escanor, Askin dan Johan Liebert.
8. Cewek-cewek RDIB; Dita, Pro, Ratna, Efa, Elisa, Dini. Sukses ya buat kalian
9. Serta semua pihak yang telah membantu dalam pengerjaan Tugas Akhir ini yang belum mampu penulis sebutkan diatas.

Terima kasih atas segala bantuan, dukungan, serta doanya. Semoga Allah SWT senantiasa memberikan kesehatan, keselamatan, karunia dan nikmat-Nya.

Penulis pun ingin memohon maaf karena Penulis menyadari bahwa Tugas Akhir ini masih belum sempurna dengan segala kekurangan di dalamnya. Selain itu penulis bersedia menerima kritik dan saran terkait dengan Tugas Akhir ini. Semoga Tugas Akhir ini dapat bermanfaat bagi seluruh pembaca.

Surabaya, Januari 2017

DAFTAR ISI

ABSTRAK	iii
ABSTRACT	v
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xvii
1. BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan	4
1.5 Manfaat	4
1.6 Relevansi	5
2. BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Sebelumnya	7
2.2 Dasar Teori	9
2.2.1 Data Cuaca Surabaya	9
2.2.1 Persiapan Data	10
2.2.2 Metode Data Mining	13
2.2.3 Matlab	30
2.2.4 Pengukuran Performa	31
3. BAB III METODOLOGI PENELITIAN	33
3.1 Studi Literatur	34
3.2 Persiapan Data	34
3.3 Klasifikasi dengan Teknik CART	34
3.4 Klasifikasi dengan Teknik Random Forest	36
3.5 Konstruksi Aplikasi	37
3.6 Pembahasan dan Dokumentasi	37
4. BAB IV PERANCANGAN	39
4.1 Statistika Deskriptif	39
4.1.1 Curah Hujan	39
4.1.2 Suhu Udara	42
4.1.3 Suhu Udara Minimum	43
4.1.4 Suhu Udara Maksimum	44
4.1.5 Titik Embun	45

4.1.6 Tekanan Pemukaan Laut.....	47
4.1.7 Visibilitas.....	48
4.1.8 Kecepatan Angin.....	49
4.1.9 Kecepatan Angin Maksimum	50
4.2 Persiapan Data.....	52
4.3 CART	52
4.3.1 Model CART Maksimal	52
4.3.2 Model dengan Pruning Optimal hasil Model CART Maksimal.....	55
4.3.3 Model dengan Pengendalian Kedalaman Pohon Keputusan	55
4.3.4 Pengujian Model CART dengan Proporsi Presentase Data Training dan Testing.....	55
4.4 Random Forest	56
4.4.1 Pemilihan Atribut Model	56
4.4.2 Model dengan Pengendalian Jumlah Pohon dan Jumlah Variabel Acak.....	58
4.4.3 Pengujian Model Random Forest dengan Proporsi Presentase Data Training dan Testing	59
4.5 Konstruksi Perangkat Lunak	59
4.5.1 Use Case Diagram	59
4.5.2 Deskripsi Use Case	60
4.5.3 Desain Antarmuka Aplikasi.....	68
5. BAB V IMPLEMENTASI	71
5.1 Data Masukan.....	71
5.2 Persiapan Data.....	71
5.2.1 Penghapusan Data Variabel Dependen.....	71
5.2.2 Multiple Imputation	72
5.2.3 Transformasi Data.....	73
5.2.4 Penurunan Fitur Baru.....	74
5.3 Implementasi Script	76
5.3.1 Memasukkan Data	77
5.3.2 CART.....	77
5.3.3 Random Forest.....	80
5.3.4 Uji Performa	82
5.4 Aplikasi Akhir.....	83
5.4.1 Fungsi Load Data.....	83

5.4.2 Fungsi Memrediksi Data dengan Metode CART Biasa	84
5.4.3 Fungsi Memrediksi Data dengan Metode CART Pruned	85
5.4.4 Fungsi Melihat Variable Importance CART	86
5.4.5 Fungsi Melihat Grafik Tree untuk CART	87
5.4.6 Fungsi Memrediksi Data dengan Metode Random Forest	87
5.4.7 Fungsi Melihat Variable Importance Random Forest	88
6. BAB VI HASIL DAN PEMBAHASAN	91
6.1 CART	91
6.1.1 Model CART Maksimal	91
6.1.2 Model dengan Pruning hasil Model CART Maksimal	102
6.1.3 Model dengan Pengendalian Kedalaman Pohon	104
6.1.4 Pengujian Model CART dengan Proporsi Presentase Data Training dan Testing	111
6.1.5 Kesimpulan Hasil Model CART	112
6.2 Random Forest	114
6.2.1 Pemilihan Atribut Model	114
6.2.2 Model dengan Pengendalian Jumlah Pohon dan Jumlah Variabel Acak	123
6.2.3 Pengujian Model Random Forest dengan Proporsi Presentase Data Training dan Testing	131
6.2.4 Kesimpulan Hasil Model Random Forest	132
6.3 Analisa Perbandingan Hasil CART dan Random Forest	134
7. BAB VII KESIMPULAN DAN SARAN	141
7.1 Kesimpulan	141
7.2 Saran	142
8. DAFTAR PUSTAKA	143
BIODATA PENULIS	149
A. LAMPIRAN A DATA MENTAH (8 VARIABEL AWAL)	A-1

B.	LAMPIRAN B DATA VARIABEL TURUNAN, CURAH HUJAN DAN KELAS INTENSITAS CURAH HUJAN.....	B-1
C.	LAMPIRAN C DATA VARIABEL TURUNAN, CURAH HUJAN DAN KELAS INTENSITAS CURAH HUJAN.....	C-1
D.	LAMPIRAN D HASIL UJI PERFORMA MODEL RANDOM FOREST	D-1

DAFTAR GAMBAR

Gambar 2-1 Struktur Simpul Pohon Keputusan.....	15
Gambar 2-2 Struktur simpul terkait <i>goodness of split</i>	17
Gambar 2-3 Hasil Pohon Keputusan.....	24
Gambar 2-4 Algoritma sederhana Random Forest.....	25
Gambar 2-5 Flowchart Algoritma Random Forest.....	26
Gambar 2-6 Confusion matrix multi kelas.....	32
Gambar 4-1 Grafik batang curah hujan.....	40
Gambar 4-2 Proporsi kejadian hujan (hujan dan tidak hujan).....	41
Gambar 4-3 Proporsi kelas intensitas hujan.....	41
Gambar 4-4 Distribusi kelas untuk suhu udara.....	42
Gambar 4-5 Distribusi kelas untuk suhu udara minimum.....	43
Gambar 4-6 Distribusi kelas untuk suhu udara maksimum.....	44
Gambar 4-7 Distribusi kelas untuk titik embun.....	46
Gambar 4-8 Distribusi kelas untuk tekanan permukaan laut.....	47
Gambar 4-9 Distribusi kelas untuk visibilitas.....	48
Gambar 4-10 Distribusi kelas untuk kecepatan angin.....	49
Gambar 4-11 Distribusi kelas untuk kecepatan angin maksimum.....	51
Gambar 4-12 Use Case Diagram Aplikasi.....	60
Gambar 4-13 Halaman memasukkan data.....	68
Gambar 4-14 Halaman pengolahan dengan teknik CART.....	69
Gambar 4-15 Halaman pengolahan dengan teknik Random Forest.....	70
Gambar 5-1 Gambar ringkasan jumlah data yang diganti nilainya.....	73
Gambar 5-2 Distribusi kelas untuk selisih suhu udara dan titik embun.....	74
Gambar 5-3 Distribusi kelas untuk estimasi kelembapan relatif.....	75
Gambar 5-4 Distribusi kelas untuk selisih suhu udara minimum dan maksimum.....	76
Gambar 5-5 Distribusi kelas untuk bulan.....	76
Gambar 5-6 Script loading data.....	77
Gambar 5-7 Script prediksi dengan CART maksimal.....	78

Gambar 5-8 Kalkulasi level pemangkasan optimal berdasarkan minimal complexity.....	78
Gambar 5-9 Script prediksi dengan CART pruned	79
Gambar 5-10 Script untuk menghasilkan grafik variable importance	80
Gambar 5-11 Script untuk menghasilkan grafik tree.....	80
Gambar 5-12 Script untuk menghasilkan prediksi dengan random forest.....	81
Gambar 5-13 Script untuk menghasilkan grafik variable importance untuk random forest.....	82
Gambar 5-14 Script untuk menghasilkan hasil uji performa ..	83
Gambar 5-15 Tampilan fungsi load data	84
Gambar 5-16 Tampilan fungsi predksi dengan metode CART	85
Gambar 5-17 Tampilan fungsi prediksi dengan CART pruned	85
Gambar 5-18 Tampilan fungsi variabel importance CART ...	86
Gambar 5-19 Tampilan fungsi grafik tree untuk CART	87
Gambar 5-20 Tampilan fungsi prediksi data dengan metode random forest.....	88
Gambar 5-21 Tampilan fungsi melihat variable importance dengan random forest	89
Gambar 6-1 Decision tree Model dengan 8 variabel	92
Gambar 6-2 Grafik kepentingan variable untuk metode CART dengan 8 variabel.....	92
Gambar 6-3 Grafik tree untuk model CART dengan 12 variabel	96
Gambar 6-4 Grafik kepentingan variabel untuk metode CART dengan 12 variabel.....	97
Gambar 6-5 Pohon keputusan hasil CART dengan pemangkasan	103
Gambar 6-6 Grafik perbandingan jumlah Minleaf (x) terhadap Minparentsiz (y) CART untuk akurasi training	105
Gambar 6-7 Grafik perbandingan jumlah Minparentsiz (x) terhadap Minleaf (y) CART untuk akurasi training	106
Gambar 6-8 Grafik perbandingan jumlah Minleaf (x) terhadap Minparentsiz (y) CART untuk akurasi testing.....	107

Gambar 6-9 Grafik perbandingan jumlah Minparentsiz (x) terhadap Minleaf (y) CART untuk akurasi testing.....	108
Gambar 6-10 Grafik perbandingan jumlah Minleaf (x) terhadap Minparentsiz (y) CART untuk Weighted Accuracy Testing.....	108
Gambar 6-11 Grafik perbandingan jumlah variabel acak (x) terhadap jumlah pohon (y) Random Forest untuk akurasi training	124
Gambar 6-12 Grafik perbandingan jumlah jumlah pohon (x) terhadap variabel acak (y) Random Forest untuk akurasi training	124
Gambar 6-13 Grafik perbandingan jumlah variabel acak (x) terhadap jumlah pohon (y) Random Forest untuk data testing	125
Gambar 6-14 Grafik perbandingan jumlah jumlah pohon (x) terhadap variabel acak (y) Random Forest untuk data testing	126
Gambar 6-15) Grafik perbandingan jumlah variabel acak (x) terhadap jumlah pohon (y) Random Forest untuk nilai weighted accuracy data testing	127
Gambar 6-16 Grafik perbandingan jumlah jumlah pohon (x) terhadap variabel acak (y) Random Forest untuk nilai weighted accuracy data testing	128

DAFTAR TABEL

Tabel 2-1 Penelitian Terkait	8
Tabel 2-2 Atribut Dataset	10
Tabel 2-3 Diskretisasi Variabel Kontinyu	13
Tabel 2-4 Data Pengamatan.....	20
Tabel 2-5 Hasil goodness of split X1 dengan median 25	21
Tabel 2-6 Hasil goodness of split X1 dengan median 35	22
Tabel 2-7 Hasil goodness of split X2 (1&2).....	22
Tabel 2-8 Hasil perbandingan <i>goodness of split</i> antar variabel dan <i>threshold</i>	23
Tabel 2-9 Hasil prediksi dengan <i>majority vote</i>	23
Tabel 2-10 Data Awal	29
Tabel 2-11 Data resample.....	29
Tabel 2-12 Hasil Prediksi dengan <i>Majority Vote</i>	30
Tabel 3-1 Metodologi Penelitian	33
Tabel 4-1 Statistika deskriptif variabel curah hujan	39
Tabel 4-2 Statistika deskriptif variabel suhu udara	42
Tabel 4-3 Statistika deskriptif variabel suhu udara minimum	43
Tabel 4-4 Statistika deskriptif variabel suhu udara maksimum.....	45
Tabel 4-5 Statistika deskriptif variabel titik embun.....	46
Tabel 4-6 Statistika deskriptif variabel tekanan permukaan laut....	47
Tabel 4-7 Statistika deskriptif variabel visibilitas	48
Tabel 4-8 Statistika deskriptif variabel kecepatan angin	50
Tabel 4-9 Statistika deskriptif variabel kecepatan angin maksimum	51
Tabel 4-10 8 Variabel awal untuk pemodelan dengan CART	53
Tabel 4-11 8 Variabel awal dan 4 variabel baru untuk pemodelan dengan CART	54
Tabel 4-12 8 Variabel awal untuk pemodelan dengan random forest	56
Tabel 4-13 8 Variabel awal dan 4 variabel baru untuk pemodelan dengan random forest	57
Tabel 4-14 Use Case Decription "Memasukkan Data"	60
Tabel 4-15 Use case description "Memrediksi Intensitas Hujan dengan Metode CART Maksimal"	61
Tabel 4-16 Use case description "Memrediksi Data dengan Metode CART Pruned"	63
Tabel 4-17 Use case description "Melihat Variable Importance CART"	64

Tabel 4-18 Use case description "Melihat Grafik Tree untuk CART"	65
Tabel 4-19 Use case description "Memrediksi Data dengan Metode Random Forest "	66
Tabel 4-20 Use case description "Melihat Variable Importance Random Forest "	67
Tabel 5-1 Statistika deskriptif variabel curah hujan	72
Tabel 5-2 Konveri satuan data	73
Tabel 6-1 Confusion Matrix Model I dengan 8 variabel	93
Tabel 6-2 Nilai akurasi CART Model I (8 variabel)	93
Tabel 6-3 Nilai precision dan recall CART Model I (8 variabel)	93
Tabel 6-4 Kepentingan variabel dari Model I (8 variabel)	94
Tabel 6-5 Nilai akurasi CART Model I (dengan pemilihan atribut)	95
Tabel 6-6 Nilai precision dan recall CART Model I (dengan pemilihan atribut)	95
Tabel 6-7 Nilai akurasi CART Model I (8 vvariabel awal dan 4 variabel baru)	97
Tabel 6-8 Nilai precision dan recall CART (8 variabel awal dan 4 variabel baru)	98
Tabel 6-9 Kepentingan variabel Model II (8 variabel awal dan 4 variabel baru)	98
Tabel 6-10 Nilai akurasi CART Model II (8 variabel awal dan 4 variabel baru)	99
Tabel 6-11 Nilai precision dan recall Model II (8 variabel awal dan 4 variabel baru)	100
Tabel 6-12 Tabel perbandingan nilai akurasi Model CART maksimal	100
Tabel 6-13 Nilai precison dan recall Model CART maksimal	101
Tabel 6-14 Nilai akurasi CART Pruned	102
Tabel 6-15 Nilai precision dan recall CART Pruned	103
Tabel 6-16 Nilai presisi CART (Pengendalian kedalaman pohon) untuk kelas 3	109
Tabel 6-17 Nilai recall CART (Pengendalian kedalaman pohon) untuk kelas 3	111
Tabel 6-18 Nilai akurasi pengujian proporsi CART	112
Tabel 6-19 Nilai parameter berdasrkan uji performa pengendalian kedalaman pohon model CART	113
Tabel 6-20 Nilai akurasi Random Forest Model I (8 variabel)	115
Tabel 6-21 Nilai precision dan recall Random forest Model I (8 variabel)	115

Tabel 6-22 Kepentingan variabel Random forest Model I (8 variabel)	116
Tabel 6-23 Nilai akurasi random forest Model I (dengan pemilihan atribut)	117
Tabel 6-24 Nilai precision dan recall random forest Model I (dengan pemilihan atribut)	117
Tabel 6-25 Nilai akurasi random forest Model I (dengan pemilihan atribut)	118
Tabel 6-26 Nilai precision dan recall (dengan pemilihan atribut)	119
Tabel 6-27 Kepentingan variabel random forest Model II (8 variabel awal, 4 variabel baru)	119
Tabel 6-28 Nilai akurasi random forest Model II (dengan pemilihan atribut)	120
Tabel 6-29 Nilai precision dan recall random forest Model II (dengan pemilihan atribut)	121
Tabel 6-30 Perbandingan nilai akurasi model random forest	121
Tabel 6-31 Perbandingan rata-rata nilai precision dan recall model random forest	122
Tabel 6-32 Nilai presisi model random forest (Pengendalian jumlah pohon dan variabel acak) kelas 3	129
Tabel 6-33 Nilai recall model random forest (Pengendalian jumlah pohon dan variabel acak) kelas 3	130
Tabel 6-34 Nilai akurasi pengujian proporsi data untuk random forest	132
Tabel 6-35 Nilai parameter berdasarkan uji performa pengendalian jumlah pohon dan jumlah variabel model random forest	133
Tabel 6-36 Perbandingan nilai uji performa CART dan Random forest	134
Tabel 6-37 Perbandingan nilai precision dan recall CART dan random forest	136
Tabel 6-38 Nilai parameter berdasarkan uji performa percobaan parameter model CART dan random forest	137
Tabel 6-39 Hasil nilai akurasi percobaan proporsi data untuk CART dan random forest	139
Tabel A-1 Data mentah (8 variabel dengan satuan Indonesia)	A-1
Tabel B-1 Data Variabel Turunan, Curah Hujan Dan Kelas Intensitas Curah Hujan	B-1
Tabel C-1 Nilai akurasi Training (CART) -Pengendalian kedalaman pohon keputusan	C-1
Tabel C-2 Nilai akurasi testing (CART) -Pengendalian kedalaman pohon keputusan	C-1

Tabel C-3 Nilai Weighted Akurasi Testing (CART) -Pengendalian kedalaman pohon keputusan.....	C-2
Tabel C-4 Nilai Presisi Kelas 0 (CART)	C-3
Tabel C-5 Nilai Presisi kelas 1 (CART)	C-3
Tabel C-6 Nilai Presisi kelas 2 (CART)	C-4
Tabel C-7 Nilai Presisi kelas 4 (CART)	C-4
Tabel C-8 Nilai Recall kelas 0 (CART).....	C-5
Tabel C-9 Nilai Recall kelas 1 (CART).....	C-6
Tabel C-10 Nilai Recall kelas 2 (CART).....	C-6
Tabel C-11 Nilai Recall kelas 4 (CART).....	C-7
Tabel D-1 Akurasi Training (Random Forest) - Pengendalian parameter	D-1
Tabel D-2 Akurasi Testing (Random Forest) - Pengendalian parameter	D-1
Tabel D-3Nilai Weighted Akurasi Testing (Random Forest) - Pengendalian parameter.....	D-2
Tabel D-4 Nilai Presisi kelas 0 (random forest)	D-3
Tabel D-5 Nilai Presisi kelas 1 (Random Forest)	D-3
Tabel D-6 Nilai Presisi kelas 2 (Random Forest)	D-4
Tabel D-7 Nilai Presisi kelas 4 (Random Forest)	D-4
Tabel D-8 Nilai Recall kelas 0 (Random Forest).....	D-5
Tabel D-9 Nilai Recall kelas 1 (Random Forest).....	D-6
Tabel D-10 Nilai Recall kelas 2 (Random Forest).....	D-6
Tabel D-11 Nilai Recall kelas 4 (Random Forest).....	D-7

LAMPIRAN

LAMPIRAN A

LAMPIRAN B

Data mentah (8 variabel awal)

Data variabel turunan, curah hujan dan kelas intensitas curah hujan

LAMPIRAN C

Hasil uji performa cart

LAMPIRAN D

Hasil uji performa model random forest

Halaman ini sengaja dikosongkan

BAB I

PENDAHULUAN

Pada bab pendahuluan diuraikan proses identifikasi masalah penelitian meliputi latar belakang masalah, perumusan masalah, batasan masalah, tujuan, manfaat kegiatan tugas akhir dan relevansi pengerjaan tugas akhir. Berdasarkan uraian pada bab ini, harapannya gambaran permasalahan dan pemecahan masalah tugas akhir dapat dipahami.

1.1 Latar Belakang

Ketersediaan data cuaca dan iklim selama beberapa dekade terakhir yang dikumpulkan melalui observasi, radar maupun satelit membuka peluang besar untuk membuat suatu alat yang akurat dan efektif untuk menganalisa pengetahuan tersembunyi dalam data tersebut. Penggalan data cuaca dan iklim memiliki tantangan besar dalam metode penggalan pengetahuan dari jumlah data yang besar [1]. Jumlah data yang sangat besar tidak lagi dapat ditangani dengan metodologi tradisional. Para peneliti melihat peluang untuk melahirkan teknologi baru yang menjawab kebutuhan ini, yaitu *data mining*. Teknologi ini diaplikasikan oleh berbagai perusahaan untuk memecahkan permasalahan bisnis dengan mengacu pada analisa data, dimana secara garis besar *data mining* berfungsi mencari pengetahuan dari data dengan jumlah yang sangat besar.

Prediksi cuaca merupakan masalah yang secara ilmiah dan teknologi menjadi perhatian dan menantang karena dipengaruhi 2 faktor utama, yaitu berpengaruh pada banyak aktivitas manusia dan karena terciptanya peluang penelitian pada bidang ini akibat semakin majunya teknologi. Pengetahuan yang didapatkan dari hasil analisa data dan peramalan cuaca dapat berperan besar dalam berbagai sektor yang dipengaruhi cuaca dan iklim secara langsung, seperti

agrikultur, pariwisata dan pengelolaan air [2]. Pada sektor transportasi terutama transportasi laut dan udara membutuhkan prakiraan cuaca yang akurat secara periodik dalam hitungan menit untuk memutuskan persetujuan penerbangan dan pelayaran, hal ini dikarenakan kondisi cuaca berkontribusi besar pada keselamatan jiwa penumpang dan awak [3].

Kebutuhan prediksi curah hujan yang akurat bermanfaat padaantisipasi banjir bandang dan manajemen saluran limbah [4]. Curah hujan bukan merupakan data deret waktu yang kontinyu karena memiliki periode kering di antara musim hujan, sehingga alat intelegensi komputasional yang konvensional kurang mampu memberikan akurasi yang dapat diterima [4]. Pendekatan data mining dengan menggunakan kombinasi multidimensional association rule mining yaitu algoritma Random Forest dan CART (Classification and Regression Trees) untuk menentukan pola keterkaitan antara curah hujan dengan faktor suhu udara, kelembapan, kecepatan angin dan tekanan permukaan laut.

Dalam penelitian [5] dilakukan prediksi kejadian hujan dengan metode klasifikasi menggunakan Decision Trees dan Random Forest yang menunjukkan bahwa performa Random Forest memiliki keunggulan dari segi kemampuan prediktif dibandingkan dengan pohon keputusan tunggal. Sementara itu CART yang merupakan pohon keputusan tunggal cenderung memiliki model kurang stabil, karena perubahan sedikit saja pada data akan berpengaruh pada hasil prediksi dan cenderung overfitting [6]. Sehingga, untuk meningkatkan kestabilan dan kemampuan prediktif, diaplikasikan metode *ensemble*, salah satu metode ensemble yaitu Random Forest yang memiliki keunggulan untuk menghindari overfitting [7].

Implementasi metode CART dan Random Forest pada data cuaca diharapkan dapat menjadi salah satu metode prediktif untuk meramalkan kejadian hujan dan intensitas hujan. Data yang digunakan berasal dari stasiun pemantauan cuaca dalam

jangka waktu 17 tahun (2000-2016) diambil dari stasiun meteorologi di Juanda, Surabaya melalui website NCDC (National Climate Data Center) yang merupakan repositori data cuaca untuk seluruh stasiun meteorologi seluruh negara. Data cuaca yang digunakan meliputi data cuaca yang berkaitan dengan curah hujan seperti suhu udara rata-rata, suhu udara maksimum, titik embun, kecepatan angin rata-rata, kecepatan angin maksimum, kecepatan angin minimum, tekanan permukaan laut, dan visibilitas.

Dengan adanya penelitian tugas akhir ini diharapkan mampu membantu adanya upaya pendeteksian potensi banjir bandang dan membantu manajemen saluran limbah, manajemen pintu air, membantu proyek pembangunan dengan memberikan prediksi potensi curah hujan pada kota Surabaya. Pengembangan perangkat lunak mengimplementasikan teknik random forest dan CART untuk memprediksi kejadian hujan dan intensitas curah hujan diharapkan mampu membantu pemerintah kota Surabaya dalam mempercepat ketanggapan proses antisipasi banjir dan manajemen saluran limbah.

1.2 Rumusan Masalah

1. Bagaimana cara mengidentifikasi atribut yang berpengaruh pada hasil prediksi kejadian hujan dan intensitas hujan di kota Surabaya?
2. Bagaimana mengaplikasikan teknik random forest dan CART untuk memprediksi kejadian hujan dan intensitas hujan di kota Surabaya?
3. Bagaimana melakukan perbandingan hasil uji performa antara metode random forest dan CART untuk memprediksi kejadian hujan dan intensitas hujan di kota Surabaya?
4. Bagaimana cara membuat aplikasi perangkat lunak dengan aplikasi Matlab dengan mengaplikasikan teknik Random Forest dan CART?

1.3 Batasan Masalah

Dari permasalahan yang disebutkan di atas, batasan masalah dalam tugas akhir ini adalah :

1. Data menggunakan data harian cuaca kota Surabaya pada stasiun meteorologi Juanda, yang didapatkan melalui website NOAA
2. Perangkat lunak dikonstruksi dengan aplikasi Matlab dengan tools Treebagger, Fitctree dan GUIDE

1.4 Tujuan

1. Mengidentifikasi atribut yang berpengaruh pada hasil prediksi kejadian hujan dan intensitas hujan di kota Surabaya
2. Mengaplikasikan teknik random forest dan CART untuk memprediksi kejadian hujan dan intensitas hujan di kota Surabaya
3. Melakukan perbandingan hasil uji performa antara metode random forest dan CART untuk memprediksi kejadian hujan dan intensitas hujan di kota Surabaya
4. Mengonstruksi perangkat lunak dengan Matlab dengan mengaplikasikan teknik Random Forest dan CART

1.5 Manfaat

1. Memberikan prediksi potensi curah hujan sehingga bisa dilakukan upaya antisipasi bencana akibat hujan seperti banjir bandang, membantu memprediksi kondisi penerbangan dan membantu majaemen saluran sanitasi.

1.6 Relevansi

Tugas akhir ini diharapkan dapat digunakan sebagai alat untuk pertimbangan prediksi kejadian dan intensitas curah hujan pada kota Surabaya serta diharapkan dapat menjadi sumber pustaka untuk penelitian terkait teknik penggalian data (*data mining*) dengan teknik random forest dan CART. Tugas akhir ini berkaitan dengan Penggalian Data dan Analitika Bisnis dan Sistem Cerdas.

Halaman ini sengaja dikosongkan

BAB II

TINJAUAN PUSTAKA

Bab ini akan menjelaskan mengenai penelitian sebelumnya dan dasar teori yang dijadikan acuan atau landasan dalam pengerjaan tugas akhir ini. Landasan teori akan memberikan gambaran secara umum dari landasan penjabaran tugas akhir ini.

2.1 Penelitian Sebelumnya

Penelitian yang dijadikan acuan dalam pengerjaan tugas akhir ini terdapat pada Tabel 2-1

Pada jurnal [5], menyatakan bahwa metode Random Forest memiliki tingkat error lebih rendah dibandingkan dengan Decision Trees dengan melakukan pencarian variabel yang penting (Important Variables) terlebih dahulu dan memperkecil kemungkinan overfitting dari Decision Tree. Penyebab teknik random forest dapat memperkecil kemungkinan overfitting ditunjang oleh jurnal [7] , karena random forest bekerja dengan mengumpulkan vote hasil prediksi terbanyak dari hasil prediksi semua decision tree di dalamnya serta kemampuan teknik random forest yang mampu memberikan tingkat akurasi prediksi hujan [5] yang relatif tinggi dengan menggunakan dataset cuaca.

Pada tugas akhir ini melakukan studi komparatif antara teknik random forest dan decision trees (CART) untuk menilai performa hasil prediksi untuk data cuaca, dengan tujuan memprediksi kejadian hujan dan intensitas hujan. Dataset yang digunakan sebelum dilakukan pengaplikasian teknik random forest dan decision tree (CART) akan dilakukan transformasi data seperti pada jurnal [8], yaitu merubah beberapa atribut dataset yang bersifat kontinyu menjadi diskrit.

Tabel 2-1 Penelitian Terkait

Judul	Metode	Penulis	Hasil yang Didapatkan
A Comparative Study on Decision Tree and Random Forest Using R Tool [5]	Prediksi kejadian hujan dengan metode klasifikasi menggunakan Decision Trees dan Random Forest	Prajwala T R	Random Forest memiliki tingkat error lebih rendah dibandingkan dengan Decision Trees dengan melakukan pencarian variabel yang penting terlebih dahulu dan memperkecil kemungkinan overfitting dari Decision Tree
A Decision Tree for Weather Prediction [8]	Metode Decision Trees (CART) untuk memprediksi suhu udara	Elia Georgiana Petre	Transformasi data kontinyu menjadi nominal/diskrit , serta pemasukan attribut tahun dan bulan sebelum diolah dengan metode Decision Tree (CART). Tingkat akurasi dihasilkan relatif tinggi yaitu 83% untuk 7 kelas klasifikasi.
A Machine Learning Approach to Finding Weather	Metode random forest, Takagi-Sugeno fuzzy untuk prediksi	John K. Williams dan D. A. Ahijevych, C. J.	Random forest dapat meminimalkan risiko overfitting yang dihasilkan

Regimes and Skillful Predictor Combinations for Short-term Storm Forecasting [7]	Consolidated Storm Prediction for Aviation (CoSPA) yaitu prediksi badai untuk penerbangan	Kessinger, T. R. Saxen, M. Steiner dan S. Dettling	dari Decision Trees. Random Forest bekerja dengan mengumpulkan vote hasil prediksi terbanyak semua decision tree di dalamnya.
--	---	--	---

2.2 Dasar Teori

Pada tahap ini akan menjelaskan hasil tinjauan pustaka yang digunakan dalam penelitian tugas akhir ini.

2.2.1 Data Cuaca Surabaya

Pada penelitian ini, dataset yang digunakan adalah data cuaca kota Surabaya dengan 9 atribut seperti pada Tabel 2-2. Dataset dipilih dari repositori data cuaca NCDC (National Climatic Data Center) yang merupakan organisasi yang menyediakan rekam data cuaca secara global untuk seluruh stasiun meteorologi. Dataset sendiri merupakan data harian ringkasan cuaca pada kota Surabaya yang direkam oleh stasiun meteorologi Juanda, Surabaya.

Berikut ini adalah klasifikasi intensitas hujan menurut Badan Meteorologi, Klimatologi dan Geofisika [9]:

Hujan ringan dengan intensitas	: 0,1 - 5,0 mm/jam atau 5
- 20	mm/hari
Hujan sedang dengan intensitas	: 5,0 - 10,0 mm/jam atau
20 - 50	mm/hari
Hujan lebat dengan intensitas	: 10,0 - 20 mm/jam atau
50 - 100	mm/hari
Hujan sangat lebat dengan intensitas	: >20 mm/jam atau >100
	mm/hari

Tabel 2-2 Atribut Dataset

No	Variabel	Deskripsi
1	<i>Temperature</i>	Rata-rata suhu udara dalam derajat Celcius
2	<i>Dew Point</i>	Suhu titik embun dalam derajat Celcius, dimana uap air mulai mengembun dari udara
3	<i>Sea Level Pressure</i>	Tekanan permukaan laut dalam milibar
4	<i>Visibility</i>	Ukuran tingkat jarak pandang dalam km
5	<i>Wind Speed</i>	Rata-rata kecepatan angin dalam knots
6	<i>Max Wind Speed</i>	Kecepatan angin maksimum yang dicatat dalam satu hari dalam knots
7	<i>Max Temperature</i>	Suhu udara maksimum dalam satu hari dalam derajat Celcius
8	<i>Min Temperature</i>	Suhu udara minimum dalam satu hari dalam derajat Celcius
9	<i>Precipitation</i>	Total curah hujan dalam mm yang akan didiskretisasi sesuai klasifikasi intensitas hujan per hari dari BMKG

2.2.1 Persiapan Data

Tahap persiapan data atau data preprocessing merupakan tahap penting dalam proses data mining. Frase “Garbage In, Garbage Out” dapat diaplikasikan pada proses data mining, dengan kata lain apabila data buruk masuk untuk diolah maka akan menghasilkan hasil yang buruk juga. Hal ini dikarenakan terkadang dalam pelaksanaan penelitian, pada proses pengumpulan data sering kurang terkontrol, misalnya adanya missing values, nilai data di luar jarak nilai yang ditetapkan (bukan outlier, namun kesalahan pemasukan data), dan lainnya [10].

Data yang masih memiliki informasi redundan, tidak relevan akan cenderung menghasilkan hasil yang menyesatkan serta

mempersulit proses penemuan pengetahuan (knowledge discovery). Sehingga, untuk meningkatkan kualitas data dan hasil data mining, tahap pemrosesan data sangat penting. Ada beberapa kategori pemrosesan data, antara lain [10]:

- a. Pembersihan Data
- b. Integrasi Data
- c. Transformasi Data
- d. Reduksi Data

2.2.1.1 Metode Complete Case (Penanganan Missing Values)

Hal paling sederhana yang harus dilakukan tentang data yang hilang (*missing values*), adalah penghapusan data dengan atribut yang memiliki nilai data yang hilang tersebut. Dalam penghapusan *listwise*, setiap baris dalam kumpulan data yang berisi ketidaklengkapan dihapus sebelum analisis dan hanya kasus lengkap dianalisis.

Metode *complete case* secara implisit mengasumsikan bahwa data hilang secara acak yaitu, tidak terkait dengan data yang hilang dan diamati dari dataset dan bila asumsi ini tidak terpenuhi, metode ini akan menghasilkan estimasi parameter bias [11]. Bahkan ketika ketidaklengkapan disebabkan oleh proses yang sepenuhnya acak, menghapus data akan mengurangi kekuatan statistik [12].

2.2.1.2 Metode Multiple Imputation (Penanganan Missing Values)

Metode Imputasi adalah pengisian nilai hilang pada suatu dataset dengan menempatkan kembali nilai hilang dengan nilai tunggal dan kemudian memperlakukannya sebagaimana nilai tersebut merupakan nilai sebenarnya. Multiple imputation memiliki sejumlah manfaat sebagai suatu

pendekatan data hilang, dengan mengisi nilai hilang sebanyak m kali imputasi.

Tahapan multiple imputations antara lain [12]:

- a. Imputasi dataset sebanyak m kali, dimana m bernilai antara 3 sampai 10
- b. Analisis data pada setiap m dataset yang diimputasi
- c. Menggunakan formula aritmatika untuk mengumulkan estimasi parameter dan standar error dari setiap analisa m dataset

Dengan menganalisis dan menggabungkan hasil dari masing-masing dataset m , *multiple imputation* menghasilkan nilai dugaan yang kurang bias dan standar error yang lebih kecil daripada yang dihasilkan oleh imputasi tunggal (*single imputation*) [11].

2.2.1.3 Transformasi Data

Teknik pengklasifikasi yang berbeda memiliki sifat yang berbeda. Kinerja pengklasifikasi tergantung pada representasi masalah. Transformasi data adalah proses dimana representasi masalah diubah. Misalnya, proses diskritisasi digunakan untuk mengkonversi dataset dengan nilai kontinyu menjadi kategoris. Pengklasifikasi *Naive Bayes* menunjukkan kinerja yang lebih baik dengan dataset terdiskritisasi [13] serta proses diskritisasi meningkatkan kecepatan pembuatan tree (dalam penentuan *split point*) karena mengurangi jumlah *split point* yang mungkin dihasilkan [14]. Hal ini menunjukkan bahwa transformasi data adalah bidang penelitian yang penting dalam pembelajaran mesin.

Contoh dari proses diskritisasi adalah pada Tabel 2-3 yang menunjukkan variabel Curah Hujan yang kontinyu, diubah menjadi variabel yang bersifat kategorikal, yaitu Intensitas

untuk dimana nilai 1 untuk curah hujan berkisar $0 < x < 0,098$, nilai 2 untuk $0,098 \leq x < 0,3$ dan nilai 3 untuk $0,3 \leq x < 2$.

Tabel 2-3 Diskretisasi Variabel Kontinyu

CURAH HUJAN	INTENSITAS
0.16	2
0.04	1
0.63	3
0.12	2

Transformasi fitur juga dapat dilakukan dengan menurunkan dari fitur lain untuk membuat fitur baru (*Derived Feature*) dengan melakukan operasi matematis pada 2 atau fitur lebih untuk menghasilkan fitur baru.

2.2.2 Metode Data Mining

Data mining adalah teknik analisis data jumlah besar untuk menemukan pola berarti dan beraturan dari sekumpulan data. Tujuan *data mining* adalah untuk meningkatkan pemasaran, penjualan, dan dukungan pelanggan melalui teknik data mining [15]. Data mining dibagi menjadi beberapa kelompok, antara lain [15] :

a. Klasifikasi

Salah satu proses data mining yang paling umum, untuk memahami data kita harus mengkategorikan untuk membuat model dugaan klasifikasi dari beberapa jenis faktor yang dapat diterapkan pada data yang belum terklasifikasikan, contohnya:

- Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan

- Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.

b. Estimasi

Digunakan untuk melakukan klasifikasi nilai variabel untuk membuat model yang menghasilkan nilai dari variabel target sebagai nilai prediksi. Contohnya memperkirakan total pendapatan rumah tangga

c. Clustering

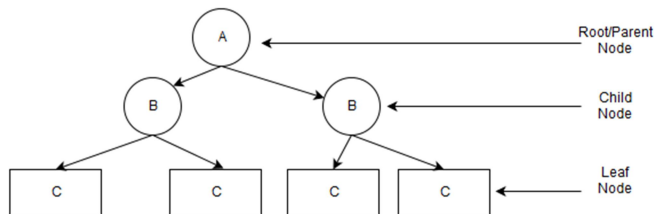
Segmentasi data heterogen menjadi berapa kelompok yang homogen, yang membedakan pengelompokan dari klasifikasi yaitu tidak bergantung pada kelas yang telah ditetapkan. Pada klasifikasi, setiap data diberikan label yang ditetapkan berdasarkan model dikembangkan melalui *preclassified.clustering* sebagai awal untuk bentuk dari data mining atau modeling. Contohnya, membagi segmentasi pelanggan ke cluster atau orang dengan kebiasaan membeli sama, untuk diketahui jenis promosi untuk tiap cluster

2.2.2.1 Decision Tree (CART)

CART (*Classification and Regression Trees*) merupakan algoritma dari teknik eksplorasi data yaitu teknik *Decision Tree* (pohon keputusan). CART menghasilkan *Classification Tree* (pohon klasifikasi) untuk nilai respon kategorik, dan menghasilkan *Regression Tree* (pohon regresi) untuk nilai responnya kontinyu. Tujuan utama CART adalah mendapatkan model yang akurat untuk melakukan pengklasifikasian. CART memiliki kelebihan yaitu merupakan metode non parametrik sehingga tidak ada asumsi variabel faktor yang harus dipenuhi, mampu mempertimbangkan interaksi antar variabel dan memudahkan dalam eksplorasi

dan pengambilan keputusan pada struktur data yang kompleks dan multi variabel (Lewis, 2000).

Struktur pohon keputusan pada CART adalah seperti Gambar 2-1. Simpul paling atas disebut simpul akar (root node), kemudian root node dibagi lagi menjadi parent node (simpul induk) yang terbagi menjadi beberapa child nodes (simpul anak) dan terakhir adalah Leaf node (simpul daun) dimana nilai kelas melekat.



Gambar 2-1 Struktur Simpul Pohon Keputusan

CART menggunakan *Gini Index* sebagai kriteria impurity. Misalkan *splitting* (pemilahan) membagi data dalam t menjadi simpul kanan tR dan simpul kiri tL , CART kemudian memilih pembagian yang memaksimalkan penurunan dalam impuritas (*decrease in impurity*). Sebagai pengganti *stopping rules* untuk menghentikan algoritma, CART menghasilkan subtree dengan menghasilkan pohon yang sangat besar kemudian memangkasnya (*pruning*) hingga simpul akhir tersisa, kemudian menggunakan *cross validation* (validasi silang) untuk memperkirakan biaya nilai misklasifikasi tiap subtree dan memilih subtree dengan nilai misklasifikasi terkecil [16]. Terdapat 3 tahap konstruksi CART, yaitu:

2.2.2.1.1 Pembuatan Pohon Keputusan Maksimal

Tahap pertama pembuatan pohon keputusan adalah bagaimana memilih *splitting attribute* (atribut untuk pemilahan dalam pembuatan pohon), tahap kedua adalah bagaimana melakukan

aturan untuk penghentian algoritma (*stopping rules*) serta tahap terakhir adalah bagaimana meletakkan simpul pada kelas [17]. Untuk lebih jelasnya, tahapan CART adalah sebagai berikut:

1. Menentukan semua kemungkinan pemilahan pada tiap atribut : Setiap splitting hasilnya tergantung pada nilai dari nilai atribut, sehingga kurang stabil karena perubahan sedikit pada data akan membuat model banyak berubah [6]. Untuk atribut yang bersifat kontinyu splitting dilakukan dengan persamaan $X_j \leq C$ dan $X_j \geq C$ (2-1):

$$X_j \leq C \text{ dan } X_j \geq C \text{ (2-1)}$$

Dimana :

$X_j = \text{Atribut ke } j$

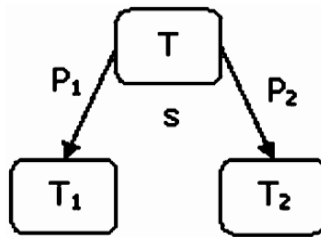
$C =$

Nilai tengah antara dua nilai amatan atribut X_j
Sehingga, jika atribut X_j
Pada atribut kategorik

Metode yang umum digunakan

Dalam Pendekatan Greedy

Atribut pemilah dipilih berdasarkan pada Goodness of Split (nilai kebaikan pemilahan). Jika kita mendefinisikan fungsi *impurity* $I(t)$ di mana t adalah setiap impuls yang diberikan, maka *Goodness of Split* didefinisikan sebagai penurunan impuritas yang dihasilkan dari pemilahan. Pada Gambar 2-2, menunjukkan kandidat pemilahan yang akan menghasilkan simpul anak T_1 dan T_2 dari T .



Gambar 2-2 Struktur simpul terkait *goodness of split*

Goodness of split merupakan perbedaan antara impurity simpul anak (T1 dan T2). Tujuannya adalah mencapai pemilahan dengan reduksi impurity paling besar. *Goodness of split* didefinisikan dengan persamaan $\Delta I(S, T) = I(t) - P_1 I(t_1) - P_2 I(t_2)$ (2-2):

$$\Delta I(S, T) = I(t) - P_1 I(t_1) - P_2 I(t_2) \quad (2-2)$$

P_1 dan P_2 = Proporsi data dari t

$I(t)$ = Fungsi *Impurity*

Fungsi *impurity* dapat didefinisikan menggunakan konsep probabilitas bersyarat $p(j|t)$. Jika ada kelas j dalam semua simpul, probabilitas kondisional $p(j|t)$ adalah probabilitas memiliki j kelas di simpul t . Probabilitas bersyarat ini probabilitas adalah N_j/N . Di mana N adalah jumlah total data di simpul dan N_j adalah jumlah kasus kelas j di simpul. Ketidakmurnian simpul adalah fungsi dari probabilitas bersyarat. CART menggunakan Gini Index untuk mendefinisikan $= i \neq j p i t p(j|t)$ (2-3):

$$l(t) = \sum_{i \neq j} p(i|t)p(j|t) \quad (2-3)$$

$p(i | t)$ =proporsi kelas i terhadap

$p(j | t)$ =proporsi kelas j terhadap

Pengukuran ini akan

2. Mendefinisikan aturan pemberhentian algoritma dimana CART akan membangun pohon hingga tidak mungkin bisa dibangun lagi, sehingga aturan pemberhentiannya adalah ketika hanya tersisa 1 data(instance) tersisa pada simpul terminal atau jika semua data di simpul hanya memiliki satu kelas.
3. Menetapkan Kelas ke Tree Nodes: Setiap simpul di pohon disertai dengan klasifikasi tertentu. Klasifikasi biasanya ditentukan oleh aturan mayoritas. Dalam sebuah simpul, kelas yang melekat akan menjadi kelas yang paling baik direpresentasikan oleh data di simpul tersebut. Berbeda dengan *Leaf node*, dimana klasifikasinya bersifat final, dan dari mereka kinerja prediksi model ditentukan.

Setiap node akan memiliki

2.2.2.1.2 Pemangkasan Pohon

Pruning (pemangkasan) dibutuhkan untuk membuat model yang lebih kecil/tidak terlalu kompleks namun tetap memiliki performa yang baik pada data baru, tidak hanya pada data uji, serta untuk menghapus simpul daun (*leaf nodes*) dengan tingkat error tinggi. Pruning ditentukan dengan *cost complexity minimum* dengan persamaan $R\alpha = R(T) + \alpha |\tilde{T}|$ (2-4) :

$$R_{\alpha} = R(T) + \alpha |\check{T}| \quad (2-4)$$

Dengan :

α =Parameter Complexity (Cost bagi penambahan satu node terminal pada pohon T)

$R(T)$

= *Penduga pengganti (resubstitution estimate) pohon T*

$|\check{T}|$ = *Ukuran banyaknya node terminal pada pohon T*

Cost complexity pruning untuk menentukan satu pohon bagian $T(\alpha)$ yang meminimumkan $R(T)$ pada seluruh pohon bagian atau untuk setiap nilai α . Kemudian, dicari pohon bagian $T(\alpha) < T_{\max}$ (pohon maksimal) yang dapat meminimumkan $R_{\alpha}(T)$ dengan tahap sebagai berikut:

1. Ambil t_R dan t_L dari T_{\max} yang dihasilkan simpul induk t
2. Bila $R(t) = R(t_R) + R(t_L)$ maka simpul t_L dan t_R dipangkas, sehingga didapatkan pohon T yang memenuhi $R(T_1) = R(T_{\max})$
3. Proses 2 diulang hingga pohon tidak bisa dipangkas.

2.2.2.1.3 Pembuatan Pohon Keputusan Optimal

Pohon optimal yang dipilih adalah pohon yang berukuran tepat dan mempunyai nilai penduga pengganti yang cukup kecil. Terdapat 2 macam penduga pengganti ($R(T)$) untuk mendapatkan pohon optimal, yaitu [18]:

- a. Penduga pengganti sampel uji ($R^{ts}(T_i)$)

$$R^{TS}(T_t) = \frac{1}{N_2} \sum_{(x_n, j_n) \in L_2} X(d(x_n) \neq j_n) \quad (2-5)$$

N_2 adalah Jumlah pengamatan dalam L_2 . Fungsi $X()$ menyatakan bila salah bernilai 0 jika benar bernilai 1. Pohon optimum dipilih dengan $R^{TS}(T_t)$ paling kecil.

- b. Penduga pengganti validasi silang v-fold ($R^{cv}(T_t)$)
v-fold berarti data pengamatan L dibagi menjadi v bagian dengan ukuran sama. Berikut ini persamaan untuk penduga pengganti untuk tiap bagian (v).

$$R(T_t^{(v)}) = \frac{1}{N_v} \sum_{(x_n, j_n) \in L_v} X(d^{(v)}(x_n) \neq j_n) \quad (2-6)$$

Dengan $N_v \cong N/V$ dimana N_v adalah jumlah pengamatan dalam N_v . Lalu setelah didapatkan penduga pengganti per bagian (v) lalu dilakukan prosedur yang sama dengan pengamatan dalam L . Penduga validasi silang untuk $T_t^{(v)}$:

$$R^{cv}(T_t) = \frac{1}{V} \sum_{v=1}^V R^{cv}(T_t^{(v)})$$

Untuk mencari pohon optimal, dicari nilai penduga pengganti validasi silang yang paling minimum.

2.2.2.1.4 Contoh CART

Berikut ini adalah contoh pembuatan pohon keputusan CART maksimal dengan menggunakan data dengan 2 variabel dependen yaitu X_1 dan X_2 dan variabel independen Y . variabel X_1 memiliki sifat kontinyu, sedangkan variabel X_2 memiliki sifat kategorikal. Berikut ini pada Tabel 2-4 adalah 20 data pengamatan

Tabel 2-4 Data Pengamatan

X_1	X_2	Y	X_1	X_2	Y	X_1	X_2	Y
20	2	1	20	1	1	30	2	2
30	2	1	20	1	1	30	2	2
30	2	1	20	1	2	30	2	2

20	1	1	40	1	2	40	2	2
20	1	1	20	1	2	40	2	2
20	2	1	40	1	2	40	2	2
40	2	1	40	1	2			

- a. Penentuan threshold dan penghitungan goodness of spilt untuk variabel X1

Penentuan threshold untuk variabel X1 yang merupakan variabel kontinyu yaitu dari nilai median antara 2 nilai unik yang berurutan, dimana terdapat n-1 atau 3-1 yaitu 2 cara pemilahan :

Median antara 20 dan 30 = 25

Median antara 30 dan 40 = 35

Kemudian, kita hitung *goodness of split* untuk 25 (Tabel 2-5):

Tabel 2-5 Hasil goodness of split X1 dengan median 25

tL : data dengan X1 = 20		tR : data dengan X1 = 30 dan 40	
NL = 8	PL = 0,4	NL = 12	PL = 0,6
N(1 t) = 6	P(1 tL) = 0,75	N(1 t) = 3	P(1 tL) = 0,25
N(2 t) = 2	P(2 tL) = 0,25	N(2 t) = 9	P(2 tL) = 0,75
i(tL) = 0,75. 0,25 + 0,25.0,75 = 0,0352		i(tL) = 0,25. 0,75 + 0,75.0,25 = 0,0352	

Sebelum mencari *goodness of split*, harus mencari nilai index gini keseluruhan:

$$N(t) = 20$$

$$N_i(t) = 9 \rightarrow p(i|t) = 9/20 = 0,45$$

$$N_j(t) = 11 \rightarrow p(j|t) = 11/20 = 0,55$$

$$\begin{aligned} i(t) &= \sum p(i|t) p(j|t) \\ &= p(1|t).p(2|t) + p(2|t).p(1|t) \\ &= 0,45.0,55 + 0,55.0,45 \\ &= 0,0613 \end{aligned}$$

Lalu memasukkan ke dalam persamaan *goodness of split* :

$$\begin{aligned} \Delta I(S, T) &= i(t) - P_1 I(t_1) - P_2 I(t_2) \\ &= 0,0613 - 0.4 (0,0352) - 0.6 (0,0352) \\ &= 0,0613 - 0,01408 - 0,02112 \end{aligned}$$

$$= 0,0261$$

Kemudian, kita hitung *goodness of split* untuk 35 (Tabel 2-6):

Tabel 2-6 Hasil goodness of split X1 dengan median 35

tL : data dengan X1 = 20 dan 30		tR : data dengan X1 = 40	
NL = 13	PL = 0,65	NL = 7	PL = 0,35
N(1 t) = 8	P(1 tL) = 0,615	N(1 t) = 1	P(1 tL) = 0,143
N(2 t) = 5	P(2 tL) = 0,385	N(2 t) = 6	P(2 tL) = 0,857
i(tL) = 0,615. 0,385 + 0,385.0,615 = 0,0561		i(tL) = 0,143. 0,857 + 0,857.0,143 = 0,0150	

Lalu memasukkan ke dalam persamaan *goodness of split* :

$$\begin{aligned}
 \Delta I(S, T) &= i(t) - P_1 I(t_1) - P_2 I(t_2) \\
 &= 0,0613 - 0,65 (0,0561) - 0,35 (0,0150) \\
 &= 0,0613 - 0,03647 - 0,00525 \\
 &= 0,01958
 \end{aligned}$$

- b. Penentuan threshold dan penghitungan goodness of split untuk variabel X2

Penentuan threshold untuk variabel X2 yang merupakan variabel kategorikal yaitu (1 dan 2), dimana terdapat $2^{(n-1)} - 1$ atau 2-1 yaitu 1 cara pemilahan : 1&2

Kemudian, kita hitung *goodness of split* untuk 1&2 (Tabel 2-7):

Tabel 2-7 Hasil goodness of split X2 (1&2)

tL : data dengan X2 = 1		tR : data dengan X2 = 2	
NL = 9	PL = 0,45	NL = 11	PL = 0,55
N(1 t) = 4	P(1 tL) = 0,44	N(1 t) = 5	P(1 tL) = 0,454
N(2 t) = 5	P(2 tL) = 0,66	N(2 t) = 6	P(2 tL) = 0,545
i(tL) = 0,44. 0,66 + 0,66.0,44		i(tL) = 0,454. 0,545 +	

= 0,0843	0,454.0,545 = 0,0612
----------	----------------------

Sebelum mencari *goodness of split*, harus mencari nilai index gini keseluruhan:

Lalu memasukkan ke dalam persamaan *goodness of split* :

$$\begin{aligned}
 \Delta I(S, T) &= i(t) - P_1 I(t_1) - P_2 I(t_2) \\
 &= 0,0613 - 0,45 (0,0843) - 0,55 (0,0612) \\
 &= 0,0613 - 0,0379 - 0,03366 \\
 &= -0,01026
 \end{aligned}$$

c. Penentuan pemilah terbaik

Kemudian kita pilih pemilah terbaik, dari hasil *goodness of split* tiap kandidat pemilah seperti pada Tabel 2-8:

Tabel 2-8 Hasil perbandingan *goodness of split* antar variabel dan *threshold*

Variabel	Threshold	Goodness of Split
X1	25	0,02610
	35	0,01958
X2	1&2	-0,01026

Terlihat dari hasil di atas, variabel X1 dengan pemilah median 25 memiliki *goodness of split* paling baik.

d. Aturan Mayoritas

Pada aturan mayoritas, kita pilih lebih banyak proporsi hasil pemilahan mana pada tiap terminal node. Seperti pada

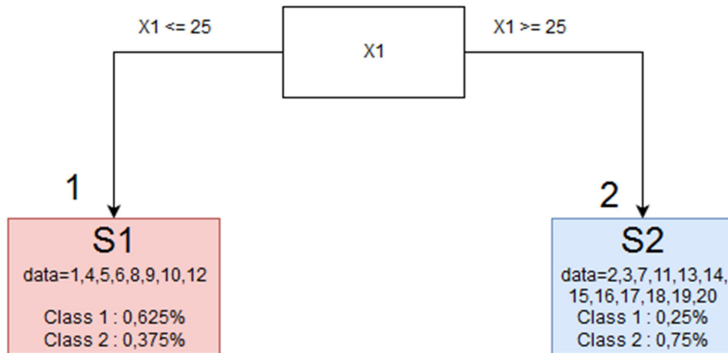
Tabel 2-9, misalnya kita bisa melihat jumlah proporsi hasil pemilahan yang memiliki kelas 1, kemudian yang memiliki proporsi kelas 1 paling banyak diberi label kelas 1.

Tabel 2-9 Hasil prediksi dengan *majority vote*

Terminal Node	N(1 t)	N(t)	p(1 t)	Label
---------------	--------	------	--------	-------

S1	5	8	0,625	1
S2	3	12	0,25	2

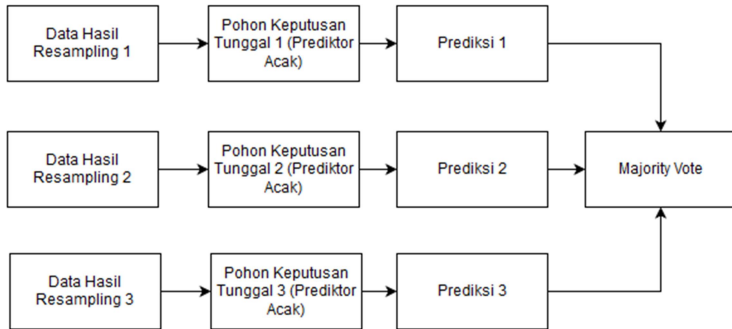
Hasil akhir pohon keputusan (Gambar 2-3) :



Gambar 2-3 Hasil Pohon Keputusan

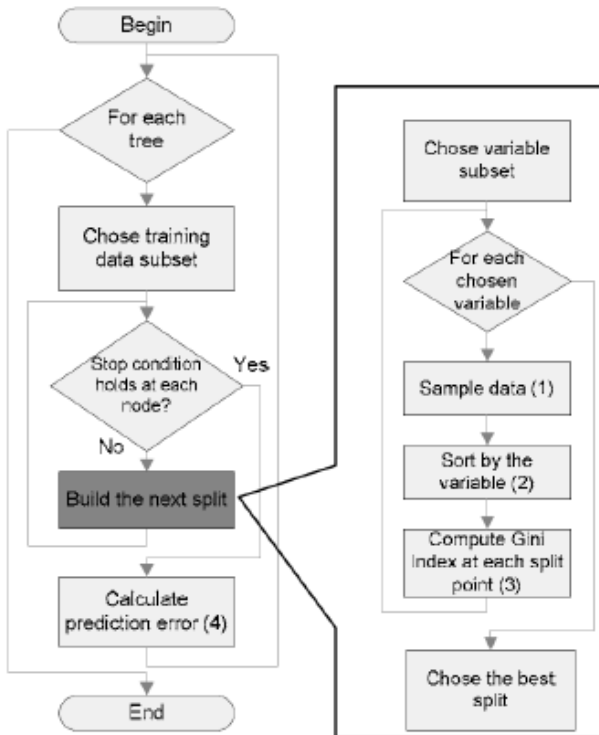
2.2.2.2 Random Forest

Random Forest (RF) adalah metode klasifikasi dan regresi berdasarkan agregasi sejumlah pohon keputusan. Secara khusus, merupakan sebuah *ensemble* (kumpulan) dari pohon yang dibangun dari kumpulan data pelatihan dan yang divalidasi secara internal untuk menghasilkan nilai respon yang dari nilai prediktor [19].



Gambar 2-4 Algoritma sederhana Random Forest

Algoritma RF secara umum dapat digambarkan seperti pada Gambar 2-4. Dalam metode RF orisinal yang disarankan oleh [17], masing-masing pohon adalah *Classification and Regresion Trees* (CART) yang menggunakan apa yang disebut *Decrease Gini Impurity* (DGI) atau penurunan impuritas Gini sebagai kriteria pemisahan dan melakukan pemilihan atribut prediktor pemisah dari subset yang dipilih secara acak dari prediktor (subset berbeda pada setiap pemisahan). Setiap pohon dibangun dari sampel bootstrap diambil dengan penggantian dari dataset asli, dan hasil prediksi semua pohon akhirnya dikumpulkan melalui *majority voting* (suara terbanyak) [17].



Gambar 2-5 Flowchart Algoritma Random Forest

Pada Gambar 2-5 merupakan alur algoritma RF, yaitu:

1. Misalkan N adalah jumlah data latih dan M adalah jumlah fitur serta m adalah jumlah fitur yang akan digunakan untuk menentukan keputusan pemilahan pada sebuah simpul, dimana m nilainya lebih kecil dari M
2. Pilih satu set latih acak dengan memilih N kali dengan pengganti dari semua data N , ini disebut juga sampel bootstrap, hanya $2/3$ dari data asli yang digunakan. Sisa dataset digunakan untuk pengujian tujuan juga disebut *Out-of-Bag* yang digunakan untuk memperkirakan error OOB untuk klasifikasi

3. Untuk setiap simpul dari pohon, simpul dipisah dengan fitur terbaik di antara fitur yang terpilih secara acak menggunakan kriteria pemisah *impurity*, lalu hitung pemisahan terbaik berdasarkan m fitur dalam data latih.
4. Setiap pohon tumbuh penuh dengan metode seperti pada subbab 2.2.2.1.1 Pembuatan Pohon Keputusan Maksimal tanpa adanya pemangkasan. Untuk prediksi, sampel yang baru, dimasukkan ke pohon, lalu diberi label/kelas sampel latih di simpul tempat sebuah sampel baru tersebut berada. Prosedur ini diulangi untuk semua pohon di dalam kumpulan *forest*, lalu hasil prediksi semua pohon diaplikasikan *majority vote* untuk diambil sebagai hasil prediksi.

Keuntungan dari RF sendiri antara lain [20] untuk menyelesaikan permasalahan multi kelas untuk klasifikasi, menghindari *overfitting*, dapat menangani nilai respon kategorikal maupun kontinyu, dapat mengetahui interaksi antar variabel.

2.2.2.2.1 Variable Importance (VIM)

Variable importance (VIM) merupakan tingkat kepentingan sebuah variabel. Terdapat 2 jenis VIM, yaitu Gini VIM dan Permutasi VIM [19]. Gini VIM adalah nilai kriteria *Decrease Gini Impurity* (DGI), diskalakan dengan total jumlah pohon di *forest*. Sebaliknya, permutasi VIM didasarkan pada akurasi prediksi bukan pada kriteria pemisahan.

Terdapat bias yang berdampak pada pemilihan kepentingan variabel, terutama pada Gini VIM. Bias ini terjadi berhubungan dengan banyaknya jumlah pemisahan kandidat pada fitur, sebagaimana dijelaskan pada bagian CART, untuk respon kategori nominal terdapat $(2^{(L-1)})-1$ kemungkinan pemisahan dan untuk respon kategori ordinal terdapat $L-1$ kemungkinan, dengan L adalah ukuran sample. Semakin banyak kandidat pemisah, semakin besar kemungkinan bahwa

paling tidak ada 1 kandidat pemisah yang memiliki kriteria pemisah yang baik, sehingga RF memilih fitur dengan banyak kelas kategori lebih sering dibandingkan fitur dengan kategori lebih sedikit. Pada permutasi VIM tidak memiliki bias, hal ini karena didasarkan pada penurunan akurasi yang dihasilkan dari hasil permutasi observasi OOB.

2.2.2.2.2 Parameter Random Forest

Melakukan *tuning* (penyetelan) parameter untuk pemodelan dengan Random Forest bertujuan untuk meningkatkan kemampuan prediksi atau membuat lebih mudah untuk melakukan pelatihan pada model. Berikut ini adalah beberapa parameter yang dapat dilakukan *tuning* [19].

- a. Jumlah Pohon
Jumlah pohon dalam *forest* akan meningkat seiring pertambahan fitur, sehingga setiap fitur memiliki cukup peluang untuk dipilih. Direkomendasikan untuk melakukan percobaan dengan beberapa nilai jumlah pohon dan berhenti hingga nilai error stabil. Meskipun begitu jumlah pohon yang lebih sedikit dapat menghasilkan akurasi sama dengan jumlah pohon lebih besar namun menghasilkan VIM yang kurang reliabel [21].
- b. Jumlah Kandidat Fitur
Jumlah kandidat fitur pada setiap pemisahan merupakan parameter dimana nilai optimalnya tergantung pada dataset. Pada metode pemilihan jumlah fitur yang acak yang disarankan Breiman [17] yaitu \sqrt{p} untuk klasifikasi dan $p/3$ untuk regresi dengan p adalah jumlah fitur, mungkin terlalu kecil terutama bila terdapat fitur dengan *noise* tinggi.

2.2.2.2.3 Contoh Random Forest

Berikut ini adalah contoh data awal (Tabel 2-10) dan data hasil resample (Tabel 2-11) sebelum diolah dengan teknik RF :

Tabel 2-10 Data Awal

Data ke	X1	X10	Y
1	5	10	1
2	1	15	1
3	5	20	1
4	1	10	1
5	5	20	2
6	1	20	2
7	5	15	2
8	5	10	2
9	5	10	2
:	:	:	:
K	5	20	1

Tabel 2-11 Data resample

Data ke	X1	X2	X9	Y
1	5	30	1	1
2	1	30	2	2
3	5	40	1	2
4	1	30	1	2
5	5	50	2	1
6	1	50	1	1
7	5	60	1	1
8	5	30	1	2
9	5	90	2	2

Inti dari algoritma RF adalah bagaimana mengambil data resample dan variabel faktor secara acak, misalnya untuk contoh ini digunakan variabel faktor sebanyak 3 ($m=3$) seperti pada tabel yaitu X1, X2 dan X9.

Tabel 2-12 Hasil Prediksi dengan *Majority Vote*

Dat a Ke	X1...	X1 0	Y	Prediksi 1	Prediksi 2	Prediksi 3	Prediks i Akhir
1	5	10	1	1	1	1	1
2	1	15	1	2	2	1	2
3	5	20	1	2	1	1	1
4	1	10	1	2	2	2	2
5	5	20	2	2	2	2	2
6	1	20	2	1	2	2	2
7	5	15	2	2	2	2	2
8	5	10	2	2	2	1	2
9	5	10	2	1	2	2	2

Hasil akhir prediksi pada

Tabel 2-12 diambil dengan aturan mayoritas semua pohon keputusan. Tingkat misklasifikasi prediksi yaitu 2/10 atau 0,2 dimana data ke 2 dan ke 5 salah diklasifikasi.

2.2.3 Matlab

Pada bagian ini akan menjelaskan teknologi yang digunakan untuk membangun perangkat lunak untuk mengakomodasi metode klasifikasi untuk prediksi kejadian curah hujan dan intensitas curah hujan, meliputi aplikasi Matlab.

Matlab memiliki kapabilitas mengintegrasikan komputasi, visualisasi, dan pemrograman. Sehingga, matlab digunakan dalam riset-riset yang memerlukan komputasi numerik kompleks. Penggunaan Matlab meliputi: matematika dan komputasi, akusisi data, pembentukan algoritma, , Analisa data, eksplorasi, dan visualisasi, pemodelan, simulasi, dan pembuatan prototype

Aplikasi matlab memfasilitasi konstruksi GUI menggunakan GUIDE serta tools yang mengakomodasi metode teknik data

mining yang digunakan pada tugas akhir ini. *Tools* Fitctree digunakan untuk pengolahan data dengan teknik CART [22], sedangkan *tools* Treebagger untuk teknik Random Forest [23].

2.2.4 Pengukuran Performa

Melakukan pengukuran terhadap metode data mining yang diterapkan pada dataset, pengukuran yang dilakukan berupa pengukuran dengan *confusion matrix* dan *cost matrix*.

2.2.4.1 Confusion Matrix

Confusion matrix merupakan sebuah Tabel untuk mendeskripsikan performa dari sebuah model klasifikasi pada sebuah data, digunakan untuk menilai kualitas output dari pengklasifikasi pada dataset [24]. Pengukuran kinerja seperti True Positive (TP), False Positive (FP), True Negative (TN), dan False Negative (FN) yang dikalkulasi untuk menilai sensitivitas, spesifisitas dan akurasi dengan persamaan:

Pada Gambar 2-6, misalnya untuk klasifikasi kelas 0, maka TP adalah pada pemetaan kelas aktual 0 ke kelas prediksi 0. FP adalah pemetaan kelas aktual non 0 ke kelas prediksi 0, TN adalah pemetaan kelas aktual non 0 yang benar diprediksikan sedangkan FN adalah pemetaan kelas aktual non 0 ke kelas prediksi non 0 yang salah.

Weighted Accuracy bertujuan untuk melihat rata-rata rasio data yang diklasifikasi dengan benar, dimana tiap kelas diberikan *weight* (bobot) yang sama untuk hasil klasifikasinya.

Confusion Matrix		Prediksi				
		0	1	2	...	j
Aktual	0	TP	FN	FN	FN	FN
	1	FP	TN	FN	FN	FN
	2	FP	FN	TN	FN	FN
	:	FP	FN	FN	TN	FN

	j	FP	FN	FN	FN	TN
--	----------	----	----	----	----	----

Gambar 2-6 Confusion matrix multi kelas

$$Recall (TP Rate) = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

$$Weighted Accuracy = \sum (\frac{TP}{TP + FN}) / Jumlah Kelas$$

BAB III

METODOLOGI PENELITIAN

Pada bab metode penelitian akan dijelaskan tahapan dilakukan dalam pengerjaan tugas akhir beserta deskripsi tiap tahapan. Pada sub bab ini akan menjelaskan mengenai metodologi pelaksanaan tugas akhir pada Tabel 3-1 :

Tabel 3-1 Metodologi Penelitian

INPUT	AKTIVITAS	OUTPUT
-Jurnal Ilmiah -Tugas Akhir/ Tesis	<div style="border: 1px solid black; padding: 5px; display: inline-block;">Studi Literatur</div> ↓	-Metode Data Mining yang cocok untuk jenis data -Metode pengukuran performa
-Dataset cuaca kota Surabaya	<div style="border: 1px solid black; padding: 5px; display: inline-block;">Persiapan Data</div> ↓	-Dataset cuaca ditransformasikan sesuai kebutuhan penelitian
-Dataset siap diolah	<div style="border: 1px solid black; padding: 5px; display: inline-block;">Klasifikasi dengan Teknik CART</div> ↓	-Hasil klasifikasi dengan CART : Pohon klasifikasi maksimal dan optimal -Hasil pengukuran performa
-Dataset siap diolah	<div style="border: 1px solid black; padding: 5px; display: inline-block;">Klasifikasi dengan Random Forest</div> ↓	-Hasil klasifikasi dengan RF, hasil tuning parameter -Hasil pengukuran performa
Model klasifikasi Random Forest, CART	<div style="border: 1px solid black; padding: 5px; display: inline-block;">Konstruksi Aplikasi</div> ↓	Konstruksi Aplikasi dengan Matlab dengan tools fitctree dan treebagger
Produk tugas akhir	<div style="border: 1px solid black; padding: 5px; display: inline-block;">Pembahasan dan Dokumentasi</div>	Laporan tugas akhir

3.1 Studi Literatur

Pada tahap ini dilakukan pengumpulan literatur yang mendukung dalam menyelesaikan tugas akhir ini. Literatur terkait penggunaan metode Random Forest dan CART serta metode pengukuran performa untuk membandingkan hasil dari kedua teknik.

3.2 Persiapan Data

Pada tahap ini akan dilakukan pemrosesan data awal sebelum data siap diolah dengan metode Random Forest dan CART, konversi satuan data, penanganan missing value pada data dengan melakukan penghapusan data untuk record yang nilai responnya tidak diketahui dan mengaplikasikan Multiple Imputation (MI) untuk mengganti nilai data pada fitur prediktor. Pada tahap ini juga akan dilakukan pembuatan fitur baru yang diturunkan dari fitur yang telah ada.

3.2.1 Statistika Deskriptif

Pada tahap ini akan dilakukan analisa variabel yang digunakan dalam tugas akhir secara deskriptif, untuk melihat nilai maksimum, rata-rata, standar deviasi, jumlah nilai unik serta persebaran distribusi data .

3.2.2 Penanganan Missing Data

Pada tahap ini, akan dilakukan penghapusan data dengan variabel depeden yang hilang serta dilakukan implementasi metode multiple imputation pada data dengan variabel independen yang hilang.

3.3 Klasifikasi dengan Teknik CART

Pada tahap ini dataset cuaca yang telah dipersiapkan sebelumnya diolah dengan metode CART menggunakan *tools* *fitctree* yang ada di aplikasi Matlab. Kemudian akan dilakukan konstruksi pohon keputusan maksimal, pemangkasan pohon keputusan dan konstruksi pohon keputusan optimal.

3.3.1 Pembuatan CART Maksimal

Pada tahap ini pohon keputusan dibiarkan untuk dibuat hingga memenuhi kriteria *stopping rules*, antara lain hingga tidak mungkin bisa dibangun lagi, sehingga aturan pemberhentiannya adalah ketika hanya tersisa 1 data (instance) tersisa pada simpul terminal atau jika semua data di simpul hanya memiliki satu kelas. Model pohon keputusan maksimal cenderung kompleks (sangat besar).

3.3.2 Variable Importance

Pada tahap ini, kepentingan variabel hasil pembuatan model yang menunjukkan kontribusi variabel pada pembentukan model. *Variable importance* ini didasarkan pada penghitungan index gini yang menunjukkan nilai *impurity* pada tiap node pada tiap pohon keputusan. Semakin sebuah variabel berkontribusi, maka semakin besar nilai kepentingan variabel.

3.3.3 Pemangkasan CART

Pada tahap ini *Pruning* (pemangkasan) dilakukan untuk membuat model yang lebih kecil/tidak terlalu kompleks namun tetap memiliki performa yang baik pada data baru, tidak hanya pada data uji, serta untuk menghapus simpul daun (*leaf nodes*) dengan tingkat error tinggi. *Pruning* ditentukan dengan *cost complexity* yang paling minimum.

3.3.4 Pembuatan CART Optimal

Pada tahap ini dibuat pohon optimal yang dipilih adalah pohon yang berukuran tepat dan mempunyai nilai penduga pengganti (*resubstitution estimate*) yang cukup kecil dengan menggunakan validasi silang.

3.3.5 Uji Skenario Parameter dan Hasil Pengukuran Performa

Pada tahap ini dilakukan pengujian skenario dengan melakukan perubahan parameter untuk membuat model CART, antara lain *Maxnumsplit*, *Minleafsize* dan

MinParentSize yang menentukan level kedalaman pembuatan pohon keputusan. Pada tiap skenario, kemudian diukur hasil performa akurasi.

3.4 Klasifikasi dengan Teknik Random Forest

Pada tahap ini akan dilakukan pengolahan data dengan teknik random forest menggunakan TreeBagger pada aplikasi Matlab. Kemudian akan dilakukan pengujian parameter untuk menilai apakah terdapat perubahan pada hasil kinerja prediktif model, antara lain seperti proporsi data latih dan uji, jumlah pohon, jumlah atribut yang dipilih acak. Kemudian akan dilakukan pengukuran performa untuk tiap metode.

3.4.1 Random Forest parameter default

Pada tahap ini akan dilakukan pembuatan model dan pengukuran hasil performa menggunakan nilai parameter yang digunakan oleh Breiman yaitu untuk sampel acak dengan proporsi $2/3$ dan jumlah prediktor acak sejumlah akar kuadrat jumlah prediktor.

3.4.2 Variable Importance

Pada tahap ini, kepentingan variabel hasil pembuatan model Random Forest menunjukkan kontribusi variabel pada pembentukan model. *Variable importance* ini didasarkan pada penghitungan index gini yang menunjukkan nilai *impurity* pada tiap node pada tiap pohon keputusan. Semakin sebuah variabel berkontribusi, maka semakin besar nilai kepentingan variabel.

3.4.3 Uji Skenario Parameter dan Hasil Pengukuran Performa

Pada tahap ini dilakukan pengujian skenario dengan melakukan pengubahan parameter untuk membuat model Random Forest, antara lain jumlah pohon dan jumlah prediktor acak yang menentukan sample yang digunakan untuk

pembuatan tiap pohon keputusan. Pada tiap skenario, kemudian diukur hasil performa akurasi.

3.5 Konstruksi Aplikasi

Tahap ini merupakan tahap konstruksi aplikasi yaitu pembangunan GUI dengan mengaplikasikan baris kode yang digunakan untuk pelaksanaan tahap klasifikasi dengan Random Forest dan CART (3.1.3). Antarmuka grafis yang dibangun akan menggunakan GUIDE dari Matlab dan *tools* *fictree* untuk membangun CART dan *tools* *treebagger* untuk Random Forest. Fungsi utama aplikasi ini adalah membuat model CART, model Random Forest dengan menggunakan data historis lalu melakukan prediksi intensitas hujan dengan menggunakan data hari sebelumnya.

3.5.1 Use Case Diagram

Pada tahap ini dibuat use case diagram untuk masing-masing fungsi yang akan diimplementasikan pada aplikasi yang akan dibuat.

3.5.2 Desain Antarmuka Aplikasi

Pada tahap ini akan dibuat desain antarmuka untuk aplikasi yang dibuat. Antarmuka meliputi input, proses serta output yang dibutuhkan untuk fungsi yang akan diimplementasikan.

3.5.3 Fungsi Aplikasi Akhir

Pada tahap ini dilakukan pengujian hasil aplikasi baik dari segi use case maupun segi hasil prediksi untuk tiap masing-masing fungsi yang diimplementasikan.

3.6 Pembahasan dan Dokumentasi

Semua proses di atas, baik pada tahapan data mining, maupun pengembangan perangkat lunak, hasil pelaksanaan akan dilakukan analisa dan pembahasannya. Kemudian, terakhir akan dilakukan dokumentasi hasil tugas akhir ini yang meliputi semua proses di atas yang sekaligus menjadi laporan dari proses pengerjaan tugas akhir yang telah dilakukan.

Halaman ini sengaja dikosongkan

BAB IV

PERANCANGAN

Bab ini menjelaskan tentang rancangan penelitian tugas akhir untuk membuat model peramalan. Bab ini berisikan proses pengumpulan data, praproses data, pengolahan data, serta bagaimana pemodelan dilakukan.

4.1 Statistika Deskriptif

Berikut ini adalah statistika deskriptif dari tiap atribut dari dataset yang digunakan dalam tugas akhir, antara lain terdapat atribut curah hujan, suhu udara, suhu udara minimum, suhu udara maksimum, titik embun, tekanan permukaan laut, visibilitas, kecepatan angin dan kecepatan angin maksimum.

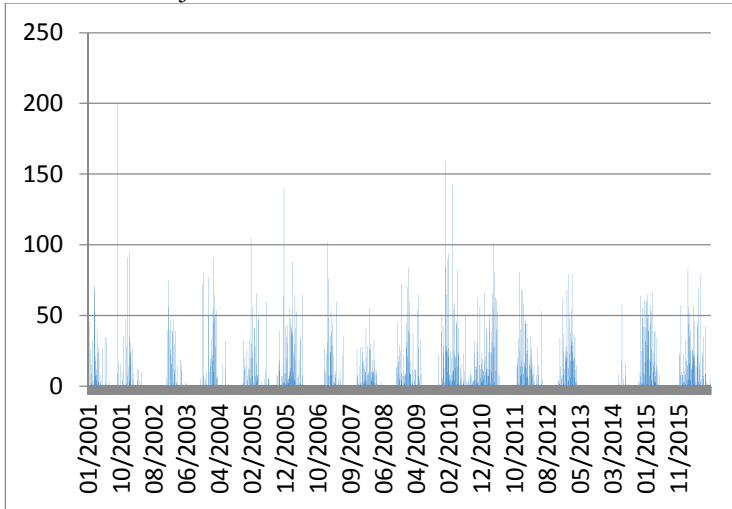
4.1.1 Curah Hujan

Pada Tabel 4-1 dapat dilihat untuk variabel curah hujan, memiliki satuan mm dengan nilai minimum 0 dan maksimum 199,9 dan rata-rata curah hujan sebesar 3,8. Terdapat 178 nilai unik, namun perlu dicatat bahwa nilai unik ini adalah hasil dari pembulatan konversi dari satuan inchi ke mm sehingga mungkin terdapat perbedaan sedikit angka desimal yang dianggap tidak memiliki perbedaan atau perbedaan nilai tidak begitu signifikan.

Tabel 4-1 Statistika deskriptif variabel curah hujan

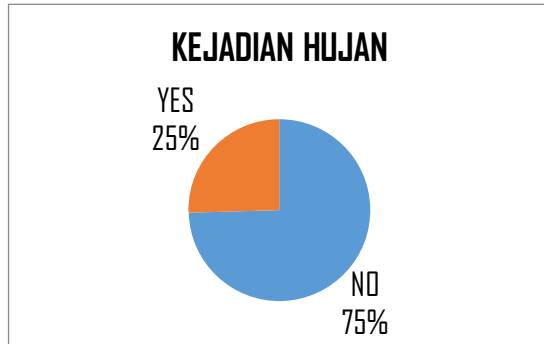
	Curah Hujan
Min	0
Max	199,9
Mean	3,825282115
Stdev	11,98499413
Nilai Unik	178
Missing	397

Dari total 6033 data, terdapat 397 data curah hujan yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong, sehingga yang digunakan dalam tugas akhir adalah sejumlah 5636 data.



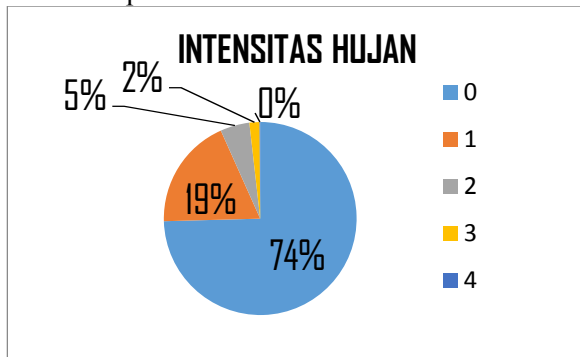
Gambar 4-1 Grafik batang curah hujan

Pada Gambar 4-1 Grafik batang curah hujan menunjukkan seasonalitas dari data curah hujan kota Surabaya yang diambil dari stasiun cuaca Juanda dari 1 Januari 2000 hingga 27 September 2016, karena terdapat jeda hari tidak hujan berturut-turut yang ditunjukkan oleh tidak adanya grafik pada jangka waktu tertentu. Hari tidak hujan berturut-turut ini dikarenakan adanya musim kemarau yang terjadi di kota Surabaya, sehingga sama sekali atau kecil kemungkinan untuk terjadi hujan.



Gambar 4-2 Proporsi kejadian hujan (hujan dan tidak hujan)

Pada Gambar 4-2 Proporsi kejadian hujan (hujan dan tidak hujan) dapat dilihat dataset yang digunakan memiliki 75% data tidak hujan atau sekitar 4202 data (curah hujan 0) dan 25% data hujan atau sekitar 1434 data (curah hujan >0). Hal ini menunjukkan adanya ketidakseimbangan data yang digunakan untuk penelitian tugas akhir dan juga sekaligus menjadi batasan penelitian.



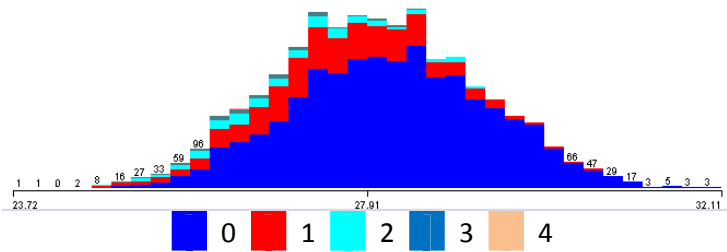
Gambar 4-3 Proporsi kelas intensitas hujan

Pada Gambar 4-3 Proporsi kelas intensitas hujan dapat dilihat, dalam penentuan kelas untuk klasifikasi, pada tugas akhir ini menggunakan 4 kelas hujan menurut BMKG [9] dan 1 kelas tidak hujan dengan proporsi 74,5% data tidak hujan (kelas

0),19% hujan ringan (keals 1), 5% hujan sedang (kelas 2), 2% hujan lebat (kelas 3) dan 0,001% hujan sangat lebat (kelas 4).

4.1.2 Suhu Udara

Pada variabel suhu udara, dapat terlihat pada Gambar 4-4 bahwa suhu udara memiliki distribusi normal, dan bila dilihat persebaran data untuk tiap kelas belum terlihat condong ke sisi mana pun (tersebar merata) atau bisa dibilang belum bisa membedakan kelas curah hujan bila melihat grafik persebaran kelas curah hujan untuk data suhu udara.



Gambar 4-4 Distribusi kelas untuk suhu udara

Tabel 4-2 Statistika deskriptif variabel suhu udara

	Suhu Udara
Min	23,72
Max	32,11
Mean	28,01138
Stdev	1,189023
Nilai Unik	131
Missing	0

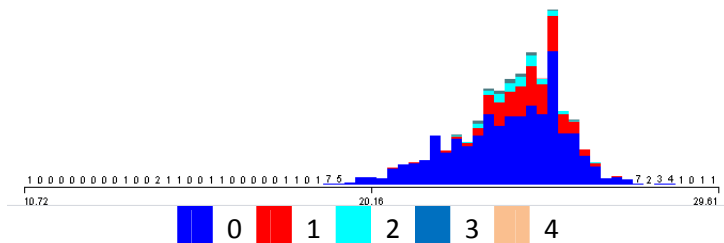
Pada Tabel 4-2 dapat dilihat untuk variabel suhu udara, memiliki satuan Celcius dengan nilai minimum 23,72 dan

maksimum 32,11 dan rata-rata suhu udara sebesar 28,011. Terdapat 131 nilai unik, namun perlu dicatat bahwa nilai unik ini adalah hasil dari pembulatan konversi dari satuan fahrenheit ke Celcius sehingga mungkin terdapat perbedaan sedikit angka desimal yang dianggap tidak memiliki perbedaan atau perbedaan nilai tidak begitu signifikan.

Dari total 5626 data, terdapat 0 data suhu udara yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong.

4.1.3 Suhu Udara Minimum

Pada variabel suhu udara minimum, dapat terlihat pada Gambar 4-5 bahwa suhu udara minimum memiliki distribusi condong ke kanan, dan bila dilihat persebaran data untuk tiap kelas belum terlihat condong ke sisi mana pun (tersebar ke sisi kanan) atau bisa dibilang belum bisa membedakan kelas curah hujan bila melihat grafik persebaran kelas curah hujan untuk data suhu udara minimum.



Gambar 4-5 Distribusi kelas untuk suhu udara minimum

Tabel 4-3 Statistika deskriptif variabel suhu udara minimum

	Suhu Udara Min.
Min	10,72
Max	29,61
Mean	23,97458659

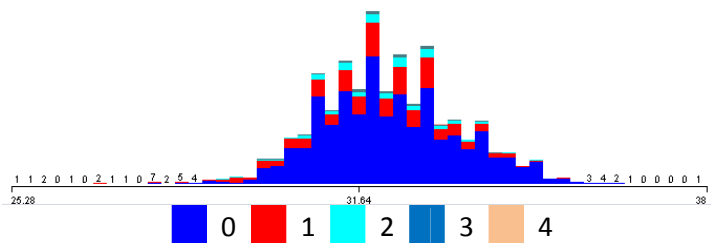
Stdev	1,477266375
Nilai Unik	107
Missing	0

Pada Tabel 4-3 dapat dilihat untuk variabel suhu udara minimum, memiliki satuan Celcius dengan nilai minimum 10,72 dan maksimum 29,61 dan rata-rata suhu udara sebesar 23,97. Terdapat 107 nilai unik, namun perlu dicatat bahwa nilai unik ini adalah hasil dari pembulatan konversi dari satuan fahrenheit ke Celcius sehingga mungkin terdapat perbedaan sedikit angka desimal yang dianggap tidak memiliki perbedaan atau perbedaan nilai tidak begitu signifikan.

Dari total 5626 data, terdapat 0 data suhu udara minimum yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong.

4.1.4 Suhu Udara Maksimum

Pada variabel suhu udara maksimum, dapat terlihat pada Gambar 4-6 bahwa suhu udara maksimum memiliki distribusi normal, dan bila dilihat persebaran data untuk tiap kelas belum terlihat condong ke sisi mana pun (tersebar merata) atau bisa dibilang belum bisa membedakan kelas curah hujan bila melihat grafik persebaran kelas curah hujan untuk data suhu udara maksimum.



Gambar 4-6 Distribusi kelas untuk suhu udara maksimum

Tabel 4-4 Statistika deskriptif variabel suhu udara maksimum

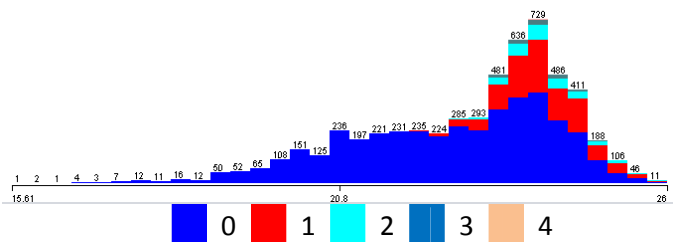
	Suhu Udara Maks.
Min	25,28
Max	38
Mean	32,17978
Stdev	1,274123
Nilai Unik	96
Missing	0

Pada Tabel 4-4 dapat dilihat untuk variabel suhu udara maksimum, memiliki satuan Celcius dengan nilai minimum 25,28 dan maksimum 38 dan rata-rata suhu udara maksimum sebesar 32,17. Terdapat 96 nilai unik, namun perlu dicatat bahwa nilai unik ini adalah hasil dari pembulatan konversi dari satuan fahrenheit ke Celcius sehingga mungkin terdapat perbedaan sedikit angka desimal yang dianggap tidak memiliki perbedaan atau perbedaan nilai tidak begitu signifikan.

Dari total 5626 data, terdapat 0 data suhu udara minimum yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong.

4.1.5 Titik Embun

Pada variabel titik embun, dapat terlihat pada Gambar 4-7 bahwa titik embun memiliki distribusi condong ke kanan, dan bila dilihat persebaran data untuk tiap kelas telah terlihat tidak tersebar merata (terlihat kelas 1 memiliki persebaran ke sisi kanan semua, sedangkan kelas 0 tersebar merata) dengan kata lain terlihat bisa membedakan kelas curah hujan bila melihat grafik persebaran kelas curah hujan untuk data titik embun.



Gambar 4-7 Distribusi kelas untuk titik embun

Tabel 4-5 Statistika deskriptif variabel titik embun

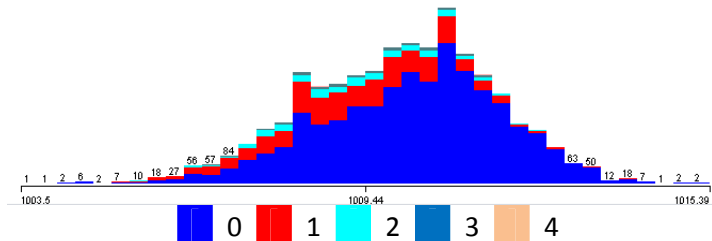
	Titik Embun
Min	15,61
Max	26
Mean	22,86587
Stdev	1,608384
Nilai Unik	164
Missing	2

Pada Tabel 4-5 dapat dilihat untuk variabel titik embun, memiliki satuan Celcius dengan nilai minimum 15,61 dan maksimum 26 dan rata-rata titik embun sebesar 22,86. Terdapat 164 nilai unik, namun perlu dicatat bahwa nilai unik ini adalah hasil dari pembulatan konversi dari satuan fahrenheit ke Celcius sehingga mungkin terdapat perbedaan sedikit angka desimal yang dianggap tidak memiliki perbedaan atau perbedaan nilai tidak begitu signifikan.

Dari total 5626 data, terdapat 2 data suhu udara minimum yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong. Data missing ini akan ditangani dengan metode Multiple Imputation (MI) sehingga tidak perlu membuang record dengan salah satu data yang hilang (missing) karena data yang hilang akan diganti dengan data hasil MI.

4.1.6 Tekanan Permukaan Laut

Pada variabel tekanan permukaan laut, dapat terlihat pada Gambar 4-8 bahwa tekanan permukaan laut memiliki distribusi normal, dan bila dilihat persebaran data untuk tiap kelas cukup terlihat tidak tersebar merata (terlihat kelas 1 memiliki persebaran ke sisi kiri, sedangkan kelas 0 tersebar merata) dengan kata lain terlihat bisa membedakan kelas curah hujan bila melihat grafik persebaran kelas curah hujan untuk tekanan permukaan laut.



Gambar 4-8 Distribusi kelas untuk tekanan permukaan laut

Tabel 4-6 Statistika deskriptif variabel tekanan permukaan laut

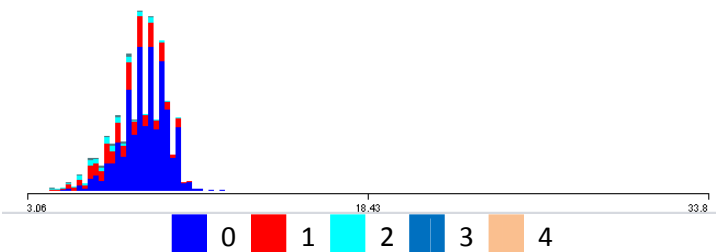
	Tekanan P. Laut
Min	1003,5
Max	1015,2
Mean	1010,019
Stdev	1,610184
Nilai Unik	106
Missing	92

Pada Tabel 4-6 dapat dilihat untuk variabel tekanan permukaan laut, memiliki satuan milibar dengan nilai minimum 1003,5 dan maksimum 1015,2 dan rata-rata titik embun sebesar 1010,019. Terdapat 106 nilai unik. Tidak terdapat konversi satuan pada data ini.

Dari total 5626 data, terdapat 92 data tekanan permukaan laut yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong. Data missing ini akan ditangani dengan metode Multiple Imputation (MI) sehingga tidak perlu membuang record dengan salah satu data yang hilang (missing) karena data yang hilang akan diganti dengan data hasil MI.

4.1.7 Visibilitas

Pada visibilitas, terlihat pada Gambar 4-9 bahwa visibilitas memiliki distribusi condong ke kiri, dan bila dilihat persebaran data untuk tiap kelas terlihat tersebar merata dengan kata lain belum bisa membedakan kelas curah hujan bila melihat grafik persebaran kelas curah hujan untuk data visibilitas.



Gambar 4-9 Distribusi kelas untuk visibilitas

Tabel 4-7 Statistika deskriptif variabel visibilitas

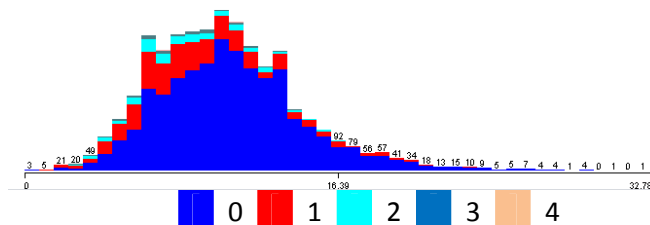
	Visibilitas
Min	3,06
Max	33,8
Mean	8,148396
Stdev	1,267564
Nilai Unik	65
Missing	0

Pada Tabel 4-7 dapat dilihat untuk variabel visibilitas, memiliki satuan km dengan nilai minimum 3,06 dan maksimum 33,8 dan rata-rata suhu udara maksimum sebesar 8,14. Terdapat 65 nilai unik, namun perlu dicatat bahwa nilai unik ini adalah hasil dari pembulatan konversi dari satuan mil ke km sehingga mungkin terdapat perbedaan sedikit angka desimal yang dianggap tidak memiliki perbedaan atau perbedaan nilai tidak begitu signifikan.

Dari total 5626 data, terdapat 0 data suhu udara minimum yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong.

4.1.8 Kecepatan Angin

Pada variabel kecepatan angin, dapat terlihat pada Gambar 4-10 bahwa kecepatan angin memiliki distribusi condong ke kiri, dan bila dilihat persebaran data untuk tiap kelas terlihat tersebar tidak merata (kelas 1 cenderung berkumpul ke sisi kiri, sementara kelas 0 menyebar merata) dengan kata lain telah bisa membedakan kelas curah hujan bila melihat grafik persebaran kelas curah hujan untuk data kecepatan angin.



Gambar 4-10 Distribusi kelas untuk kecepatan angin

Tabel 4-8 Statistika deskriptif variabel kecepatan angin

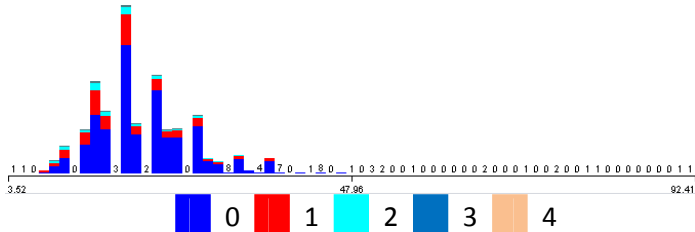
	Kecepatan Angin
Min	0
Max	32,78
Mean	10,46892
Stdev	3,926628
Nilai Unik	153
Missing	0

Pada Tabel 4-8 dapat dilihat untuk variabel kecepatan angin, memiliki satuan km/h dengan nilai minimum 0 dan maksimum 32,78 dan rata-rata kecepatan angin sebesar 10,46. Terdapat 153 nilai unik, namun perlu dicatat bahwa nilai unik ini adalah hasil dari pembulatan konversi dari satuan knots ke km/h sehingga mungkin terdapat perbedaan sedikit angka desimal yang dianggap tidak memiliki perbedaan atau perbedaan nilai tidak begitu signifikan.

Dari total 5626 data, terdapat 0 data suhu udara minimum yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong.

4.1.9 Kecepatan Angin Maksimum

Pada variabel kecepatan angin maksimum, dapat terlihat pada Gambar 4-11 bahwa kecepatan angin maksimum memiliki distribusi condong ke kiri, dan bila dilihat persebaran data untuk tiap kelas terlihat tersebar merata dengan kata lain belum bisa membedakan kelas curah hujan bila melihat grafik persebaran kelas curah hujan untuk data kecepatan angin maksimum.



Gambar 4-11 Distribusi kelas untuk kecepatan angin maksimum

Tabel 4-9 Statistika deskriptif variabel kecepatan angin maksimum

	Kecepatan Angin Maks.
Min	3,52
Max	92,41
Mean	20,65032
Stdev	6,770893
Nilai Unik	43
Missing	7

Pada Tabel 4-9 dapat dilihat untuk variabel kecepatan angin maksimum, memiliki satuan km/h dengan nilai minimum 3,52 dan maksimum 92,41 dan rata-rata kecepatan angin sebesar 20,65. Terdapat 43 nilai unik, namun perlu dicatat bahwa nilai unik ini adalah hasil dari pembulatan konversi dari satuan knots ke km/h sehingga mungkin terdapat perbedaan sedikit angka desimal yang dianggap tidak memiliki perbedaan atau perbedaan nilai tidak begitu signifikan.

Dari total 5626 data, terdapat 7 data suhu udara minimum yang memiliki nilai missing, dimana dalam dataset ditunjukkan dengan data kosong. Data missing ini akan ditangani dengan metode Multiple Imputation (MI) sehingga tidak perlu membuang record dengan salah satu data yang

hilang (missing) karena data yang hilang akan diganti dengan data hasil MI.

4.2 Persiapan Data

Pada tahap ini akan dilakukan beberapa pemrosesan pada data, antara lain:

- a. Penghapusan data curah hujan (variabel dependen) yang missing
- b. Penggantian variabel independen yang missing dengan metode Multiple Imputation dengan 5 kali iterasi menggunakan SPSS 17.0
- c. Konversi satuan data dari satuan USA ke satuan Indonesia
- d. Penurunan fitur baru meliputi Selisih Temperature dan Dew Point, Estimasi Kelembapan Relatif, Selisih Temperature Minimum dan Maksimum dan Bulan

4.3 CART

Pada tahap ini dilakukan pemodelan dengan menggunakan metode CART.

4.3.1 Model CART Maksimal

Pohon keputusan CART akan dibiarkan berkembang tanpa adanya pruning atau pemotongan dengan menggunakan 70% data sebagai data training. Pada tahap ini akan dilakukan uji coba kombinasi variabel mana saja yang akan memberikan tingkat kinerja yang baik. Tingkat kinerja akan diukur dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.3.1.1 Model I - dengan 8 Variabel Awal

Pada Model I, CART maksimal akan dibuat dengan melibatkan 8 variabel dari dataset awal, dapat dilihat pada Tabel 4-10:

Tabel 4-10 8 Variabel awal untuk pemodelan dengan CART

No	Variable	Keterangan
1.	Temp	Suhu Udara
2.	Dewp	Suhu titik embun
3.	Slp	Tekanan permukaan laut
4.	Visib	Visibilitas
5.	Wdsp	Kecepatan Angin
6.	Mxspd	Kecepatan Angin Maksimum
7.	MaxTemp	Suhu Udara Maksimum
8.	MinTemp	Suhu Udara Minimum

Kemudian diukur tingkat kinerja model dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.3.1.2 Model dengan Variabel hasil Pemilihan Variable Importance dari Model I

Model I pada tahap 4.3.1.1 Model I - dengan 8 Variabel Awaldiukur variable importance (tingkat kepentingan variabel) untuk melihat variabel mana yang berkontribusi paling besar untuk kinerja model. Pemilihan variabel dengan memilih nilai importance variabel di atas rata-rata nilai importance variabel keseluruhan.

Kemudian diukur tingkat kinerja model dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.3.1.3 Model II - dengan 8 Variabel Awal dan 4 Variabel Turunan

Pada Model I, CART maksimal akan dibuat dengan melibatkan 8 variabel (variabel 1-8) dari dataset awal dan 4 variabel (variabel 9-12) turunan, dapat dilihat pada Tabel 4-11

Tabel 4-11 8 Variabel awal dan 4 variabel baru untuk pemodelan dengan CART

No	Variable	Keterangan
1.	Temp	Suhu Udara
2.	Dewp	Suhu titik embun
3.	Slp	Tekanan permukaan laut
4.	Visib	Visibilitas
5.	Wdsp	Kecepatan Angin
6.	Mxspd	Kecepatan Angin Maksimum
7.	MaxTemp	Suhu Udara Maksimum
8.	MinTemp	Suhu Udara Minimum
9.	RTDp	Jarak suhu udara dan suhu titik embun
10.	EstRelH	Estimasi kelembapan relatif
11.	RtminTmax	Jarak suhu udara maksimum dan minimum
12.	Month	Bulan

Kemudian diukur tingkat kinerja model dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.3.1.4 Model dengan Variabel hasil Pemilihan Variable Importance dari Model II

Model I pada tahap 4.3.1.3 Model II - dengan 8 Variabel Awal dan 4 Variabel Turunan diukur variable importance (tingkat kepentingan variabel) untuk melihat variabel mana yang berkontribusi paling besar untuk kinerja model. Pemilihan variabel dengan memilih nilai importance variabel di atas rata-rata nilai importance variabel keseluruhan.

Kemudian diukur tingkat kinerja model dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.3.2 Model dengan Pruning Optimal hasil Model CART Maksimal

Dari keempat model pada tahap 4.3.1 Model CART Maksimal dipilih model dengan tingkat kinerja yang baik. Kemudian model pohon keputusan yang dibiarkan berkembang maksimal tersebut dipangkas (dipruning), pemangkasan optimal dicari dengan melakukan minimasi *resubstitution etimate* dengan *cross validation* (validasi silang).

4.3.3 Model dengan Pengendalian Kedalaman Pohon Keputusan

Pada tahap ini pohon keputusan dibuat dengan melakukan tuning parameter yang mempengaruhi tingkat kedalaman (*depth*) atau besarnya pohon keputusan. Parameter yang dilakukan tuning adalah Minleafsize (jumlah observasi pada simpul *leaf*) dan MinParent Size (jumlah observasi minimum pada simpul *parent*). Berikut ini adalah nilai default Minleafsize dan Minparentsized:

1. Minleafsize (default: 10)
2. MinParentSize (default: 1)

Model dilakukan tuning dengan nilai parameter Minleafsize (1,3,5,10,20,30,40,50,60,70,80,90,100) dan nilai parameter MinParentSize (1,3,5,10,30,50,70,90). Kemudian diukur tingkat kinerja model dengan masing-masing kombinasi parameter dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing dan dipilih model dengan tingkat kinerja yang paling baik.

4.3.4 Pengujian Model CART dengan Proporsi Presentase Data Training dan Testing

Pada tahap ini, dari 3 model yang telah dihasilkan yaitu Model CART maksimal, model CART dengan pruning dan model dengan pengendalian kedalaman pohon keputusan, dipilih 1 model yang memiliki tingkat kinerja paling baik.

Kemudian model terbaik tersebut dilakukan pengujian proporsi data training dan testing antara lain (60%:40%, 75%:25%, 80%:20%, 85%:15%, 90%:10%) untuk melihat stabilitas tingkat kinerja model.

4.4 Random Forest

Model random forest dibuat dengan menggunakan 70% data sebagai data training. Pada tahap ini akan dilakukan uji coba kombinasi variabel mana saja yang akan memberikan tingkat kinerja yang baik, tuning parameter model dan pengujian proposi data. Tingkat kinerja tiap model akan diukur dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.4.1 Pemilihan Atribut Model

Pada tahap ini dilakukan pemilihan variabel untuk diolah dengan metode CART. Terdapat 4 model, yaitu model dengan 8 variabel, model dengan pemilihan atribut (dari 8 variabel), model dengan 12 variabel dan model dengan pemilihan atribut (dari 12 variabel).

4.4.1.1 Model I - dengan 8 Variabel Awal

Pada Model I, model random forest akan dibuat dengan melibatkan 8 variabel dari dataset awal, dapat dilihat pada Tabel 4-12:

Tabel 4-12 8 Variabel awal untuk pemodelan dengan random forest

No	Variable	Keterangan
1.	Temp	Suhu Udara
2.	Dewp	Suhu titik embun
3.	Slp	Tekanan permukaan laut
4.	Visib	Visibilitas
5.	Wdsp	Kecepatan Angin
6.	Mxspd	Kecepatan Angin Maksimum
7.	MaxTemp	Suhu Udara Maksimum
8.	MinTemp	Suhu Udara Minimum

Kemudian diukur tingkat kinerja model dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.4.1.2 Model dengan Variabel hasil Pemilihan Variable Importance dari Model I

Model I pada tahap 4.4.1.1 Model I - dengan 8 Variabel Awal diukur variable importance (tingkat kepentingan variabel) untuk melihat variabel mana yang berkontribusi paling besar untuk kinerja model. Pemilihan variabel dengan memilih nilai importance variabel di atas rata-rata nilai importance variabel keseluruhan.

Kemudian diukur tingkat kinerja model dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.4.1.3 Model II - dengan 8 Variabel Awal dan 4 Variabel Turunan

Pada Model I, model random forest akan dibuat dengan melibatkan 8 variabel dari dataset awal dan 4 variabel turunan, dapat dilihat pada Tabel 4-13:

Tabel 4-13 8 Variabel awal dan 4 variabel baru untuk pemodelan dengan random forest

No	Variable	Keterangan
1.	Temp	Suhu Udara
2.	Dewp	Suhu titik embun
3.	Slp	Tekanan permukaan laut
4.	Visib	Visibilitas
5.	Wdsp	Kecepatan Angin
6.	Mxspd	Kecepatan Angin Maksimum
7.	MaxTemp	Suhu Udara Maksimum
8.	MinTemp	Suhu Udara Minimum
9.	RTDp	Jarak suhu udara dan suhu titik embun
10.	EstRelH	Estimasi kelembapan relatif
11.	RtminTmax	Jarak suhu udara maksimum

No	Variable	Keterangan
		dan minimum
12.	Month	Bulan

Kemudian diukur tingkat kinerja model dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.4.1.4 Model dengan Varibel hasil Pemilihan Variable Importance dari Model II

Model I pada tahap 4.4.1.4 Model dengan Varibel hasil Pemilihan Variable Importance dari Model II diukur variable importance (tingkat kepentingan variabel) untuk melihat variabel mana yang berkontribusi paling besar untuk kinerja model. Pemilihan variabel dengan memilih nilai importance variabel di atas rata-rata nilai importance variabel keseluruhan.

Kemudian diukur tingkat kinerja model dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing.

4.4.2 Model dengan Pengendalian Jumlah Pohon dan Jumlah Variabel Acak

Pada tahap ini pohon keputusan dibuat dengan melakukan tuning parameter dari atribut yang dipilih pada tahap 4.4.1 Pemilihan Atribut Model. Parameter yang dilakukan tuning adalah jumlah pohon dan jumlah variabel acak.

Model dilakukan tuning dengan nilai parameter jumlah pohon (10,20,30,40,50,80,100,150,200,250) dan nilai parameter jumlah variabel acak (1,2,3,4,5,6,7,8,9,10,11,12). Kemudian diukur tingkat kinerja model dengan masing-masing kombinasi parameter dengan akurasi data training, akurasi data testing dan akurasi terbobot data testing dan dipilih model dengan tingkat kinerja yang paling baik.

4.4.3 Pengujian Model Random Forest dengan Proporsi Presentase Data Training dan Testing

Pada tahap ini, dari model yang telah dihasilkan yaitu model dengan kombinasi parameter jumlah pohon dan jumlah variabel acak, dipilih 1 model yang memiliki tingkat kinerja paling baik.

Kemudian model terbaik tersebut dilakukan pengujian proporsi data training dan testing antara lain (60%:40%, 75%:25%, 80%:20%, 85%:15%, 90%:10%) untuk melihat stabilitas tingkat kinerja model.

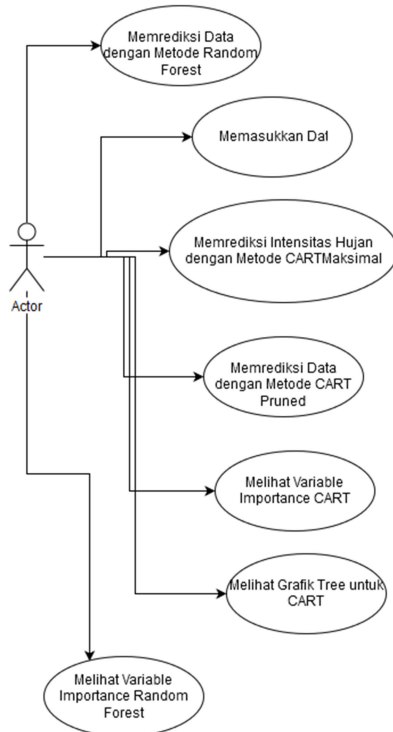
4.5 Konstruksi Perangkat Lunak

Berikut ini adalah perancangan perangkat lunak untuk aplikasi prediksi intensitas curah hujan kota Surabaya, meliputi pembuatan use case diagram dan use case description.

4.5.1 Use Case Diagram

Pada Gambar 4-12 ini adalah use case diagram untuk perancangan perangkat lunak untuk aplikasi prediksi intensitas curah hujan kota Surabaya, terdapat 7 use case utama, antara lain:

- a. Memasukkan Data
- b. Memrediksi Intensitas Hujan dengan Metode CART Maksimal
- c. Memrediksi Data dengan Metode CART Pruned
- d. Melihat Variable Importance CART
- e. Melihat Grafik Tree untuk CART
- f. Memrediksi Data dengan Metode Random Forest
- g. Melihat Variable Importance Random Forest



Gambar 4-12 Use Case Diagram Aplikasi

4.5.2 Deskripsi Use Case

Berikut ini adalah deskripsi dari 7 use case utama dalam konstruksi aplikasi.

1. Memasukkan Data

Pada

Tabel 4-14 merupakan use case description untuk “memasukkan data”

Tabel 4-14 Use Case Description "Memasukkan Data"

UC-1 Memasukkan Data	
Purpose	Memasukkan data yang akan diolah
Overview	Dimulai ketika user berada di tab Data, lalu memilih

	file excel sebagai input data	
Actors	Pengguna aplikasi	
Pre Condition	Pengguna telah membuka aplikasi	
Post Condition	Data telah masuk ke aplikasi	
Typical Course Event	Actor	System
	1. Memilih file excel yang menjadi sumber data	2. Mengambil data di excel, tiap kolom data menjadi variabel yang akan diolah
		3. Menampilkan semua nama variabel
	4. Mencentang variabel yang ingin dijadikan masukan model	6. Mengambil index variabel kemudian menyimpannya menjadi input variabel model
	5. Mengklik tombol "Select Atributtes"	7. Mengambil semua data untuk tiap variabel
		8. Menampilkan index dan nama variabel pada textbox.
Alternate Flow of Events		
Exceptional Flow of Events	Aplikasi tertutup otomatis karena error Step 3 : Nama variable tidak dapat ditampilkan karena format yang berbeda . Step 7 : Data variabel tidak dapat diambil karena format data tidak dapat terbaca.	

2. *Memrediksi Intensitas Hujan dengan Metode CART Maksimal*

Pada Tabel 4-15 merupakan use case description untuk "Memrediksi Intensitas Hujan dengan Metode CART Maksimal"

Tabel 4-15 Use case description "Memrediksi Intensitas Hujan dengan Metode CART Maksimal"

UC-2 Memrediksi Intensitas Hujan dengan Metode CART Maksimal

Purpose	Memrediksi intensitas hujan dengan Metode CART Maksimal	
Overview	Dimulai ketika user berada di tab CART, dengan data sudah terinput dan user melakukan training model lalu memprediksi intensitas hujan dengan data hari sebelumnya	
Actors	Pengguna aplikasi	
Pre Condition	-Pengguna telah membuka aplikasi -Pengguna telah memasukkan data untuk diolah -Data pada index terakhir dijadikan input untuk prediksi sesuai index atribut yang dipilih	
Post Condition	Hasil prediksi intensitas hujan dihasilkan	
Typical Course Event	Actor	System
	1. Mengklik tombol "CART"	2. Menampilkan halaman untuk pembuatan model dan prediksi dengan teknik CART
	3. Memasukkan nilai %Trainsize (0-1) 4. Memasukkan Nilai MinLeaf 5. Memasukkan Nilai MinParentSize 6. Klik tombol "START"	7. Menampilkan nilai TestSize (1-%Trainsize) 8. Menampilkan nilai level tree 9. Menampilkan hasil kinerja model (index attribut, nama attribut, hasil akurasi train, hasil akurasi tes, confusion matrix, sensitivity dan specifity)
	10. Mengklik tombol "Prediksi"	11. Menghasilkan hasil kelas prediksi dari data index terakhir
Alternate Flow of Events	Step 4: Tidak memasukkan nilai Minleaf, maka menggunakan nilai default: Step 5: Tidak memasukkan nilai MinParentsiz, maka menggunakan nilai default	
Exceptional Flow of Events	Aplikasi tertutup otomatis karena error	

3. Memrediksi Data dengan Metode CART Pruned

Pada Tabel 4-16 merupakan use case description untuk “Memrediksi Data dengan Metode CART Pruned”

Tabel 4-16 Use case description "Memrediksi Data dengan Metode CART Pruned"

UC-3 Memrediksi Data dengan Metode CART Pruned		
Purpose	Memrediksi Data dengan Metode CART Pruned (Dipangkas)	
Overview	Dimulai ketika user berada di tab CART dan user telah membuat model dengan CART maksimal lalu melakukan pruning model , lalu memprediksi intensitas hujan dengan data hari sebelumnya	
Actors	Pengguna aplikasi	
Pre Condition	-Pengguna telah membuka aplikasi -Pengguna telah memasukkan data untuk diolah -Data pada index terakhir dijadikan input untuk prediksi sesuai index atribut yang dipilih -Pengguna telah membuat model dengan CART Maksimal	
Post Condition	Hasil prediksi intensitas hujan dihasilkan	
Typical Course Event	Actor	System
	1. Mengklik tombol “Best Level” 4. Memasukkan nilai best level pada textbox 5. Mengklik tombol “PRUNE”	2. Mengkalkulasi nilai best level untuk pruning (pemangkasan) 3. Menampilkan nilai best level 6. Memangkas model sesuai nilai level pruning 7. Menampilkan hasil kinerja model yang terpangkas (index atribut, nama atribut, hasil akurasi train, hasil akurasi tes, confusion matrix,

		sensitivity dan specificity)
	8. Mengklik tombol “Prediksi”	9. Menghasilkan hasil kelas prediksi dari data index terakhir
Alternate Flow of Events	Step 1 & 2 : Tidak menghitung nilai Best Level, namun langsung ke Step 3 dengan memasukkan nilai level yang akan dipangkas langsung	
Exceptional Flow of Events	Aplikasi tertutup otomatis karena error Step 5 : memasukkan nilai yang akan dipruning namun lebih besar daripada nilai level maksimum, menghasilkan error	

4. *Melihat Variable Importance CART*

Pada

Tabel 4-17 merupakan use case description untuk “Melihat Variable Importance CART”

Tabel 4-17 Use case description "Melihat Variable Importance CART"

UC-4 Melihat Variable Importance CART		
Purpose	Melihat Variable Importance CART	
Overview	Dimulai ketika user berada di tab CART dan user telah membuat model dengan CART lalu mengklik tombol “Variable Importance” untuk melihat grafik kepentingan attribut	
Actors	Pengguna aplikasi	
Pre Condition	-Pengguna telah membuka aplikasi -Pengguna telah memasukkan data untuk diolah -Data pada index terakhir dijadikan input untuk prediksi sesuai index atribut yang dipilih -Pengguna telah membuat model dengan CART (Maksimal atau Pruned)	
Post Condition	Grafik “Variable Importance”	
Typical Course Event	Actor	System
	1. Mengklik tombol “Variable	2. Menghasilkan grafik Variabel Importance

	Importance”	
Alternate Flow of Events		
Exceptional Flow of Events	Aplikasi tertutup otomatis karena error	

5. *Melihat Grafik Tree untuk CART*

Pada

Tabel 4-18 merupakan use case description untuk “Melihat Grafik Tree untuk CART”

Tabel 4-18 Use case description "Melihat Grafik Tree untuk CART"

UC-5 Melihat Grafik Tree untuk CART		
Purpose	Melihat Grafik Tree untuk CART	
Overview	Dimulai ketika user berada di tab CART dan user telah membuat model dengan CART lalu mengklik tombol “Tree” untuk melihat grafik tree model CART (Maksimal atau Pruned)	
Actors	Pengguna aplikasi	
Pre Condition	-Pengguna telah membuka aplikasi -Pengguna telah memasukkan data untuk diolah -Data pada index terakhir dijadikan input untuk prediksi sesuai index atribut yang dipilih -Pengguna telah membuat model dengan CART (Maksimal atau Pruned)	
Post Condition	Grafik “Variable Importance”	
Typical Course Event	Actor	System
	1. Mengklik tombol “Tree”	2. Menghasilkan grafik pohon keputusan (tree) untuk model
Alternate Flow of Events		
Exceptional Flow of Events	Aplikasi tertutup otomatis karena error	

6. *Memrediksi Data dengan Metode Random Forest*

Pada

Tabel 4-19 merupakan use case description untuk “Memrediksi Data dengan Metode Random Forest”.

Tabel 4-19 Use case description "Memrediksi Data dengan Metode Random Forest "

UC-6 Memrediksi Data dengan Metode Random Forest		
Purpose	Memrediksi Data dengan Metode Random Forest	
Overview	Dimulai ketika user berada di tab RANDOMFOREST, dengan data sudah terinput dan user melakukan training model lalu memprediksi intensitas hujan dengan data hari sebelumnya	
Actors	Pengguna aplikasi	
Pre Condition	-Pengguna telah membuka aplikasi -Pengguna telah memasukkan data untuk diolah -Data pada index terakhir dijadikan input untuk prediksi sesuai index atribut yang dipilih	
Post Condition	Hasil prediksi intensitas hujan dihasilkan	
Typical Course Event	Actor	System
	1. Mengklik tombol “RANDOMFOREST”	2. Menampilkan halaman untuk pembuatan model dan prediksi dengan teknik RANDOMFOREST
	3. Memasukkan nilai %Trainsize (0-1) 4. Memasukkan Nilai Atribut Acak 5. Memasukkan Nilai Jumlah Tree 6. Klik tombol “START”	7. Menampilkan nilai TestSize (1-%Trainsize) 8. Menampilkan hasil kinerja model (index atribut, nama atribut, hasil akurasi train, hasil akurasi tes, confusion matrix, sensitivity dan specificity)
	9. Mengklik tombol “Prediksi”	10. Menghasilkan hasil kelas prediksi dari data index terakhir

Alternate Flow of Events	Step 4 : Tidak memasukkan nilai Atribut Acak, maka aplikasi akan menggunakan nilai default yaitu \sqrt{m}
Exceptional Flow of Events	Aplikasi tertutup otomatis karena error Step 4 : Memasukkan nilai Atribut acak lebih dari jumlah total atribut, maka aplikasi akan error

7. *Melihat Variable Importance Random Forest*

Pada

Tabel 4-20 merupakan use case description untuk “Melihat Variable Importance Random Forest”.

Tabel 4-20 Use case description "Melihat Variable Importance Random Forest "

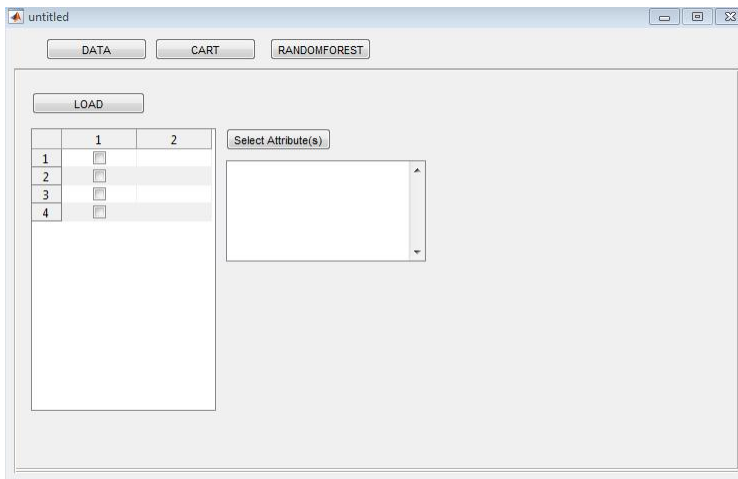
UC-7 Melihat Variable Importance Random Forest		
Purpose	Melihat Variable Importance Random Forest	
Overview	Dimulai ketika user berada di tab RANDOMFOREST dan user telah membuat model dengan RANDOMFOREST lalu mengklik tombol “Variable Importance” untuk melihat grafik kepentingan atribut	
Actors	Pengguna aplikasi	
Pre Condition	-Pengguna telah membuka aplikasi -Pengguna telah memasukkan data untuk diolah -Data pada index terakhir dijadikan input untuk prediksi sesuai index atribut yang dipilih -Pengguna telah membuat model dengan Random Forest	
Post Condition	Grafik “Variable Importance”	
Typical Course Event	Actor	System
	1. Mengklik tombol “Variable Importance”	2. Menghasilkan grafik Variabel Importance
Alternate Flow of Events		
Exceptional Flow of Events	Aplikasi tertutup otomatis karena error	

4.5.3 Desain Antarmuka Aplikasi

Berikut ini adalah penjelasan desain antarmuka aplikasi yang terdiri dari 3 halaman, yaitu halaman memasukkan data, halaman mengolah dengan teknik CART dan halaman mengolah dengan random forest.

A. Halaman Memasukkan Data

Pada Gambar 4-13 adalah halamam untuk memasukkan dan memilih atribut data yang akan diolah, terdiri dari tombol LOAD untuk memilih file excel berisi data, tabel list atribut data dan tombol “select attributes” untuk memilih mana saja atribut yang akan dipilih.

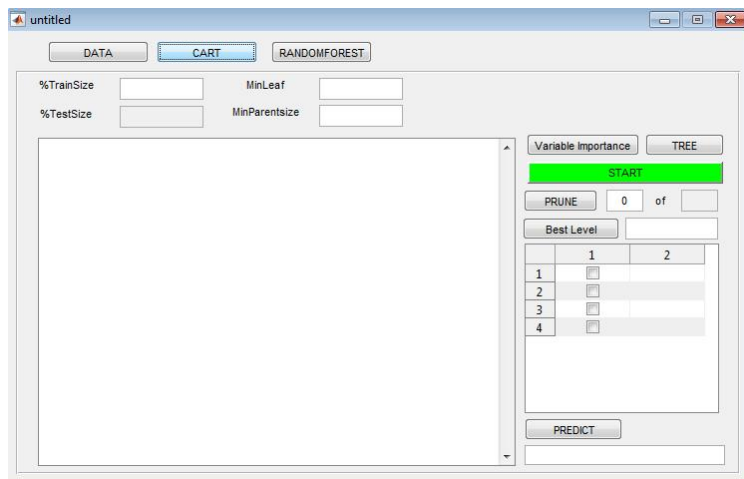


Gambar 4-13 Halaman memasukkan data

B. Halaman Pengolahan dengan Teknik CART

Pada halaman CART seperti pada Gambar 4-14, memfasilitasi seleksi proporsi data training dan testing, nilai parameter minimal leaf node dan minimal parent node, yang akan

menghasilkan confusion matrix dan hasil uji performa berupa akurasi data training, akurasi data testing dan weighted accuracy data testing. Untuk menghasilkan CART optimal, dapat dilakukan pemangkasan dengan terlebih dahulu mengkalkulasi level pemangkasan terbaik berdasarkan nilai minimal complexity. Halaman ini memfasilitasi juga pembuatan grafik tree model dan grafik tingkat kepentingan variabel, serta mampu melakukan prediksi kelas intensitas hujan menggunakan data dengan index paling akhir.

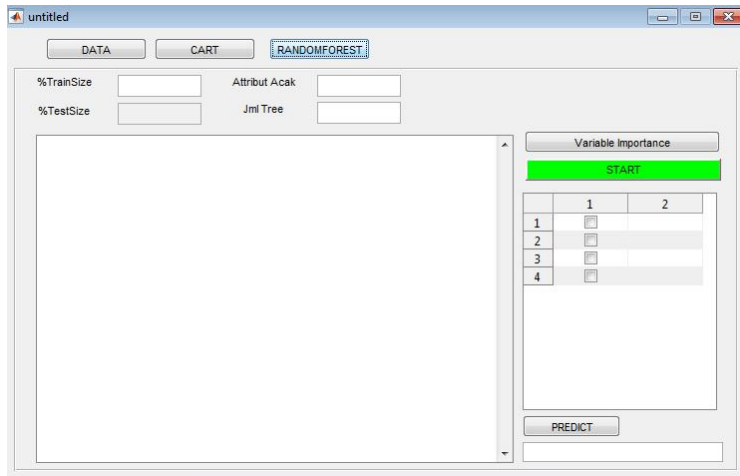


Gambar 4-14Halaman pengolahan dengan teknik CART

C. Halaman Pengolahan dengan Tenik Random Forest

Pada halaman RANDOMFOREST seperti pada Gambar 4-15, memfasilitasi seleksi proporsi data training dan testing, nilai parameter jumlah variabel acak dan jumlah tree, yang akan menghasilkan confusion matrix dan hasil uji performa berupa akurasi data training, akurasi data testing dan weighted accuracy data testing. Halaman ini memfasilitasi juga pembuatan grafik tingkat kepentingan variabel, serta mampu

melakukan prediksi kelas intensitas hujan menggunakan data dengan index paling akhir.



Gambar 4-15 Halaman pengolahan dengan teknik Random Forest

BAB V

IMPLEMENTASI

Bab ini menjelaskan proses pelaksanaan penelitian dan pembuatan model yang akan digunakan untuk pengolahan dengan teknik CART dan random forest.

5.1 Data Masukan

Data masukan untuk pembuatan model prediksi kelas intensitas curah hujan dengan teknik CART dan Random Forest menggunakan 12 variabel (8 variabel dataset asli dan 4 variabel yang diturunkan), terdiri dari suhu udara, suhu udara minimum, suhu udara maksimum, titik embun, tekanan permukaan air laut, visibilitas, kecepatan angin, kecepatan angin maksimum, selisih suhu udara dan titik embun, estimasi kelembapan relatif, selisih suhu udara minimum dan maksimum serta bulan. Proporsi data yang digunakan sebagai awal eksperimen adalah 70% data training dan 30% data testing.

5.2 Persiapan Data

Pada tahap ini akan dilakukan persiapan data (*pre-processing data*) sebelum data diolah dengan menggunakan teknik CART dan Random Forest. Data dilakukan penghapusan data variabel dependen (curah hujan), penggantian nilai hilang dengan Multiple Imputation (MI), lalu data yang masih dengan satuan USA dikonversi menjadi satuan yang digunakan di Indonesia, kemudian diturunkan beberapa fitur baru dari fitur yang ada pada dataset.

5.2.1 Penghapusan Data Variabel Dependen

Variabel dependen, atau dengan kata lain variabel yang dipengaruhi dan variabel yang diobservasi dalam penelitian ini apabila nilainya tidak ada atau *missing*, maka dilakukan penghapusan record yang mengandung missing data tersebut.

Tabel 5-1 Statistika deskriptif variabel curah hujan

	Curah Hujan
Min	0
Max	199,9
Mean	3,825282115
Stdev	11,98499413
Nilai Unik	178
Missing	397

Hal ini dikarenakan apabila yang tidak ada adalah beberapa nilai variabel independen, maka hubungan antara variabel dependen dengan variabel independen yang lain masih bisa diobservasi, namun dengan tidak adanya variabel dependen maka tidak bisa mengobservasi hubungan antara variabel independen dengan dependen, sehingga record harus dihapus. Pada Tabel 5-1 dapat dilihat total data 6033, missing 397 data curah hujan, sisanya digunakan yaitu 5636.

5.2.2 Multiple Imputation

Pada tahap ini dilakukan metode Multiple Imputation (MI) menggunakan SPSS 17.0. Pada Gambar 5-1 dapat dilihat untuk data yang hilang antara lain 2 data untuk variabel DEWP (titik embun), 92 data untuk SLP (tekanan permukaan laut) dan 7 data untuk MXSP (Kecepatan angin maksimum). Imputasi dilakukan dalam 5 iterasi, kemudian hasil iterasi kelima (iterasi terakhir) diambil untuk proses berikutnya.

Tiap kali iterasi, menghasilkan jumlah data sesuai data yang hilang sehingga terlihat pada gambar, hasil Imputed Valued (nilai imputasi) merupakan 5 kali dari jumlah data yang hilang.

Imputation Models

	Model		Missing Values	Imputed Values
	Type	Effects		
DEWP	Linear Regression	TEMP,SLP, VISIB,WDSP, MXSP,MAX, MIN,PRCP	2	10
SLP	Linear Regression	TEMP,DEWP, VISIB,WDSP, MXSP,MAX, MIN,PRCP	92	460
MXSP	Linear Regression	TEMP,DEWP, SLP,VISIB, WDSP,MAX, MIN,PRCP	7	35

Gambar 5-1 Gambar ringkasan jumlah data yang diganti nilainya

5.2.3 Transformasi Data

Pada tahap ini dilakukan konversi satuan data dari satuan USA (sebagaimana asal sumber agregator data) ke satuan Indonesia, pada Tabel 5-2 dilampirkan satuan asal (USA) dan satuan setelah data dikonversi (Indonesia):

Tabel 5-2 Konveri satuan data

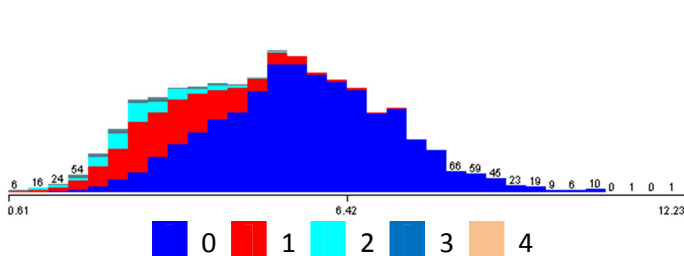
Variabel	Satuan USA	Satuan Indonesia
<i>Temperature</i>	Fahrenheit	Celcius
<i>Dew Point</i>	Fahrenheit	Celcius
<i>Sea Level Pressure</i>	hPa	hPa
<i>Visibility</i>	miles	km
<i>Wind Speed</i>	knots	km/jam
<i>Max Wind Speed</i>	knots	km/jam
<i>Max Temperature</i>	Fahrenheit	Celcius
<i>Min Temperature</i>	Fahrenheit	Celcius
<i>Precipitation</i>	inch	mm

5.2.4 Penurunan Fitur Baru

Berikut ini adalah penurunan fitur (atribut) baru dari 8 variabel awal pada dataset.

a. Selisih Temperature dan Dew Point

Fitur suhu udara yang merupakan pengukuran rata-rata suhu dalam 1 hari dengan beberapa kali observasi dan titik embun yang merupakan suhu udara dimana uap air mulai mengalami kondensasi (pendinginan/pengembunan). Pada Gambar 5-2 menunjukkan semakin kecil selisih antara suhu udara dengan titik embun, maka akan semakin tinggi intensitas hujan yang terjadi. Rumus penghitungan selisih adalah $ABS(TEMP-DEWP)$.



Gambar 5-2 Distribusi kelas untuk selisih suhu udara dan titik embun

b. Estimasi Kelembapan Relatif

Estimasi kelembapan relatif diturunkan dari suhu udara dan titik embun dengan menggunakan formula perkiraan August-Roche-Magnus. Formula tersebut mengimplikasikan bahwa tekanan saturasi uap air berubah secara eksponensial. Berikut ini adalah formula perkiraan August-Roche-Magnus untuk estimasi kelembapan relatif [25] :

$$RH = 100 \frac{\exp(\frac{aT_d}{b+T_d})}{\exp(\frac{aT}{b+T})}$$

Dimana :

a = 17,625

b = 243.04

T = suhu udara dalam Celcius

Td = titik embun dalam Celcius

RH = kelembapan relatif dalam %

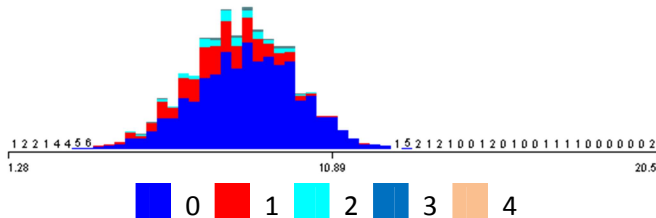
Berdasarkan perkiraan August-Roche-Magnus, estimasi kelembapan relatif tersebut valid untuk:

$0^{\circ}\text{C} < T < 60^{\circ}\text{C}$

$1\% < RH < 100\%$

$0^{\circ}\text{C} < Td < 50^{\circ}\text{C}$

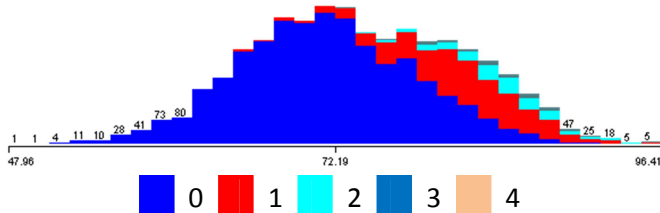
Pada Gambar 5-3 menunjukkan distribusi kelas hujan untuk estimasi kelembapan relatif.



Gambar 5-3 Distribusi kelas untuk estimasi kelembapan relatif

c. Selisih Temperature Minimum dan Maksimum

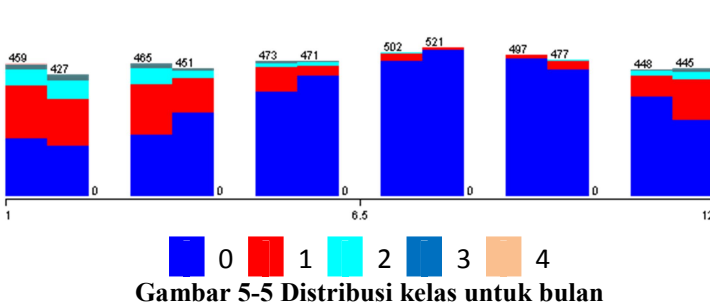
Melakukan observasi apakah terdapat hubungan signifikan antara jangka suhu antara suhu udara minimum dan suhu udara maksimum dalam 1 hari observasi terhadap intensitas curah hujan jika dibandingkan fitur yang lain. Penghitungan selisih suhu udara minimum dan maksimum adalah ABS (MAX-MIN). Pada Gambar 5-4 menunjukkan semakin besar selisih maka semakin besar intensitas hujan yang terjadi.



Gambar 5-4 Distribusi kelas untuk selisih suhu udara minimum dan maksimum

d. Bulan

Fitur bulan merupakan fitur yang diturunkan dari data timestamp data harian curah hujan. Fitur bulan dimasukkan karena sifat hujan yang seasonal atau musiman seperti pada Gambar 5-5 yang menunjukkan proporsi kelas tidak hujan membesar untuk bulan mei (5) sampai oktober (10) dan mengecil untuk bulan november (11) sampai april (4), fitur bulan sebagai atribut waktu dijadikan salah satu fitur untuk melakukan observasi apakah terdapat hubungan signifikan antara bulan dan intensitas curah hujan jika dibandingkan fitur yang lain.



Gambar 5-5 Distribusi kelas untuk bulan

5.3 Implementasi Script

Berikut ini adalah penjelasan script utama yang digunakan dalam menjalankan penelitian dan membangun aplikasi.

5.3.1 Memasukkan Data

Penjelasan mengenai script untuk memasukkan data dari file excel.

a. Load Data

```
[filename, pathname] = ...
    uigetfile({'*.xlsx'; '*.xls'}, 'File
Selector');
name = fullfile (pathname,filename)
[num,txt,row] = xlsread(name)
X = transpose (num(1:end-1,:))
Y = txt(2:end-1,end)
xpred = transpose(num(end,:))
abs = cell(numel(row (1,1:end-1)),2)
```

Gambar 5-6 Script loading data

Gambar 5-6 bertujuan untuk mengambil file excel, kemudian membaca semua data berbentuk numerik yang disimpan dalam variabel num (untuk variabel independen) serta data berbentuk string disimpan dalam bentuk txt (untuk data kelas intensitas hujan, data dependen), data yang diambil adalah semua index kecuali index terakhir yang akan dijadikan input untuk melakukan prediksi. Nama variabel dari dataset akan disimpan pada variabel abs.

5.3.2 CART

Penjelasan script untuk mengelola data dengan metode CART.

a. Prediksi dengan CART Maksimal

Pada Gambar 5-7 dapat dijelaskan pembuatan model CART dengan menggunakan fitctree dengan parameter minleaf (minimal leaf node) dan minparentsiz (minimal parent node), kemudian dilakukan perbandingan hasil klasifikasi model untuk data training (dengan membandingkan data train dengan hasil prediksi training(predTrain)) serta melakukan perbandingan hasil klasifikasi model untuk data testing. Tabel

confusion matrix untuk uji performa data training disimpan dalam variabel C, sementara untuk tabel confusion matrix untuk uji performa data testing disimpan dalam variabel D.

```
tree1 =
fitctree(transpose(Xtrain),Ytrain,'MinLeaf',leaf
n,'MinParentSize',parentn)
predTrain = predict(tree1,transpose(Xtrain))
predTest = predict(tree1,transpose(Xtest))

[C,Corder] = confusionmat(Ytrain,predTrain,...
'order',{'tidak','ringan','sedang','lebat','sang
atlebat'})

[D,Dorder] = confusionmat(Ytest,predTest,...
'order',{'tidak','ringan','sedang','lebat','sang
atlebat'})

[E,Eorder] = confusionmat(Ytest,Ytest,...
'order',{'tidak','ringan','sedang','lebat','sang
atlebat'})
```

Gambar 5-7 Script prediksi dengan CART maksimal

b. Kalkulasi level optimal CART

Pada Gambar 5-8 dapat dijelaskan ,kalkulasi level optimal, bertujuan sebagai masukan rekomendasi level tree yang dapat dipangkas dengan berdasarkan kriteria minimal complexity yang ditunjukkan oleh opsi 'min', dengan menggunakan nilai penduga pengganti silang/*cross validation* (cvloss) dengan 3 bagian (fold).

```
[E,SE,Nleaf,bestlevel] = cvLoss(tree1,...
'subtrees','all','treesize','min',
'KFold',3)
```

Gambar 5-8 Kalkulasi level pemangkasan optimal berdasarkan minimal complexity

c. Prediksi dengan CART Pruned

Pada Gambar 5-9 dapat dijelaskan, pembuatan model CART yang terpangkas dengan menggunakan fungsi `prune` yang meminta parameter 'level' yaitu level pemangkasan, dimana level tidak boleh melebihi level maksimal, kemudian dilakukan perbandingan hasil klasifikasi model untuk data training (dengan membandingkan data train dengan hasil prediksi training(`predTrain`)) serta melakukan perbandingan hasil klasifikasi model untuk data testing. Tabel confusion matrix untuk uji performa data training disimpan dalam variabel C, sementara untuk tabel confusion matrix untuk uji performa data testing disimpan dalam variabel D.

```
tree1 = prune(tree1, 'level', lvlprune)
predTrain = predict(tree1, transpose(Xtrain))
predTest = predict(tree1, transpose(Xtest))

[C, Corder] = confusionmat(Ytrain, predTrain, ...
    'order', {'tidak', 'ringan', 'sedang', 'lebat', 'sang
    atlebat'})

[D, Dorder] = confusionmat(Ytest, predTest, ...
    'order', {'tidak', 'ringan', 'sedang', 'lebat', 'sang
    atlebat'})

[E, Eorder] = confusionmat(Ytest, Ytest, ...
    'order', {'tidak', 'ringan', 'sedang', 'lebat', 'sang
    atlebat'})
```

Gambar 5-9 Script prediksi dengan CART pruned

d. Melihat Grafik Variable Importance CART

Script pada Gambar 5-10 menghasilkan grafik kepentingan variabel dengan menggunakan fungsi `predictorImportance`, tiap variabel akan direpresentasikan dalam diagram batang dengan label nama variabel yang diambil dari variabel `abs`.

```

tree1 = evalin ('base', 'tree1')
imp = predictorImportance(tree1);
abs = evalin ('base', 'abs')
figure;
bar(imp);
title('Predictor Importance Estimates');
ylabel('Estimates');
xlabel('Predictors');
index=find([abs{:},1]==true)
Labels = abs([index],2)
h = gca;
h.XTickLabel = Labels;

```

Gambar 5-10 Script untuk menghasilkan grafik variable importance

e. Melihat grafik Tree

Script pada Gambar 5-11 menghasilkan grafik tree berdasarkan model CART yang telah dibuat dengan metode fitctree. Grafik yang dihasilkan memperlihatkan level kedalaman pohon, rule di tiap node serta label kelas pada tiap leaf node.

```

tree1 = evalin ('base', 'tree1')
view(tree1,'mode','graph')

```

Gambar 5-11 Script untuk menghasilkan grafik tree

5.3.3 Random Forest

a. Prediksi dengan Random Forest

Pada Gambar 5-12 dapat dijelaskan pembuatan model Random Forest dengan menggunakan fungsi TreeBagger dengan parameter ntree (jumlah pohon) dan NVarToSample (jumlah variabel acak), kemudian dilakukan perbandingan hasil klasifikasi model untuk data training (dengan membandingkan data train dengan hasil prediksi training(predTrain)) serta melakukan perbandingan hasil klasifikasi model untuk data

testing. Tabel confusion matrix untuk uji performa data training disimpan dalam variabel C, sementara untuk tabel confusion matrix untuk uji performa data testing disimpan dalam variabel D.

```
Mdl =
TreeBagger(ntreen,transpose(Xtrain),transpose(Yt
rain),'Method','classification','oobpred','on','
oobvarimp','on',...
'NVarToSample',mtryn);
predTrain = predict(Mdl,transpose(Xtrain))
predTest = predict(Mdl,transpose(Xtest))

[C,Corder] = confusionmat(Ytrain,predTrain,...
'order',{'tidak','ringan','sedang','lebat','sang
atlebat'})

[D,Dorder] = confusionmat(Ytest,predTest,...
'order',{'tidak','ringan','sedang','lebat','sang
atlebat'})

[E,Eorder] = confusionmat(Ytest,Ytest,...
'order',{'tidak','ringan','sedang','lebat','sang
atlebat'})
```

Gambar 5-12 Script untuk menghasilkan prediksi dengan random forest

f. Melihat grafik varibale importance Random Forest

Script pada Gambar 5-13 menghasilkan grafik kepentingan variabel dengan menggunakan fnksi predictorImportance, tiap variabel akan direpresentasikan dalam diagram batang dengan label nama variabel yang diambil dari variabel abs.

```
figure;
bar(Mdl.OOBPermutedVarDeltaError);
xlabel('Feature(s)');
index=find([abs(:,1)]==true)
Labels = abs([index],2)
h = gca;
h.XTickLabel = Labels;
ylabel('Out-of-Bag Feature Importance');
title('Predictor Importance Estimates');
```

Gambar 5-13 Script untuk menghasilkan grafik variable importance untuk random forest

5.3.4 Uji Performa

Penjelasan script untuk melakukan uji performa metode CART dan Random Forest.

a. Uji Performa

Script pada Gambar 5-14, bertujuan untuk menghitung hasil pengukuran performa menggunakan hasil akurasi data training (acctrain) dan data testing (acctest) dengan menghitung jumlah diagonal (jumlah data yang berhasil diklasifikasikan benar) dibandingkan total data. Kemudian menghitung weigted accuracy untuk data testing dengan menghitung hasil rata-rata pembagian data yang diklasifikasikan benar dibandingkan total data tiap kelas.

```
AccTrain = sum(diag(C))/numel(predTrain)
AccTest = sum(diag(D))/numel(predTest)
Wacc = mean(diag(D)./diag(E))

temp = diag(D)

precision(n,m,:,1)=temp(1)/sum(D(:,1))
precision(n,m,:,2)=temp(2)/sum(D(:,2))
precision(n,m,:,3)=temp(3)/sum(D(:,3))
precision(n,m,:,4)=temp(4)/sum(D(:,4))
```

```

precision(n,m,:,5)=temp(5)/sum(D(:,5))

recall(n,m,:,1)=temp(1)/sum(D(1,:))
recall(n,m,:,2)=temp(2)/sum(D(2,:))
recall(n,m,:,3)=temp(3)/sum(D(3,:))
recall(n,m,:,4)=temp(4)/sum(D(4,:))
recall(n,m,:,5)=temp(5)/sum(D(5,:))

```

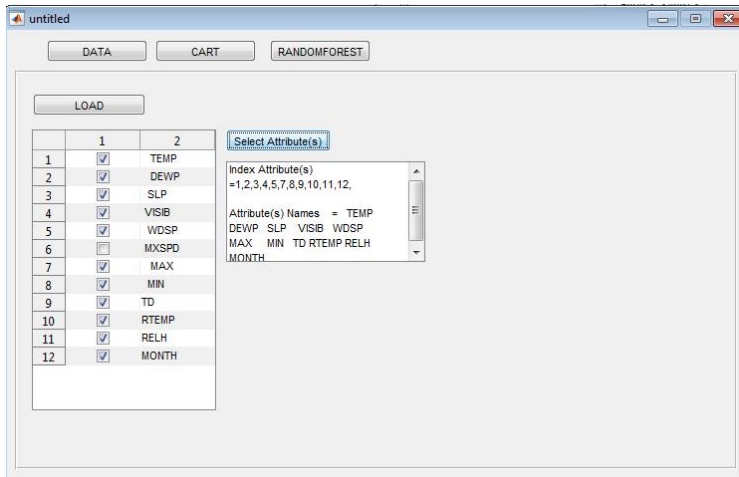
Gambar 5-14 Script untuk menghasilkan hasil uji performa

5.4 Aplikasi Akhir

Berikut ini adalah penjelasan tentang fungsi untuk aplikasi akhir.

5.4.1 Fungsi Load Data

Pada Gambar 5-15 dapat dijelaskan Load data memiliki fungsi untuk memilih file excel. Fungsi tersebut kemudian mengambil semua data dalam excel untuk dimasukkan dalam aplikasi, kemudian tiap kolom direpresentasikan sebagai variabel yang memiliki index, nama header setiap kolom dijadikan nama yang mewakili atribut. Memfasilitasi pemilihan mana saja atribut yang akan diolah, yang kemudian akan menampilkan index atribut dan nama atribut yang dipilih.

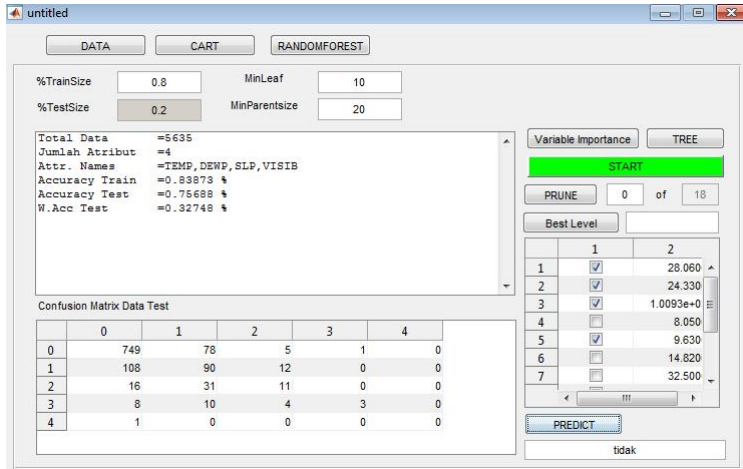


Gambar 5-15 Tampilan fungsi load data

5.4.2 Fungsi Memrediksi Data dengan Metode CART Biasa

Pada Gambar 5-16 dapat dijelaskan pembuatan model pohon keputusan menggunakan CART dengan atribut yang telah dipilih pada load data.

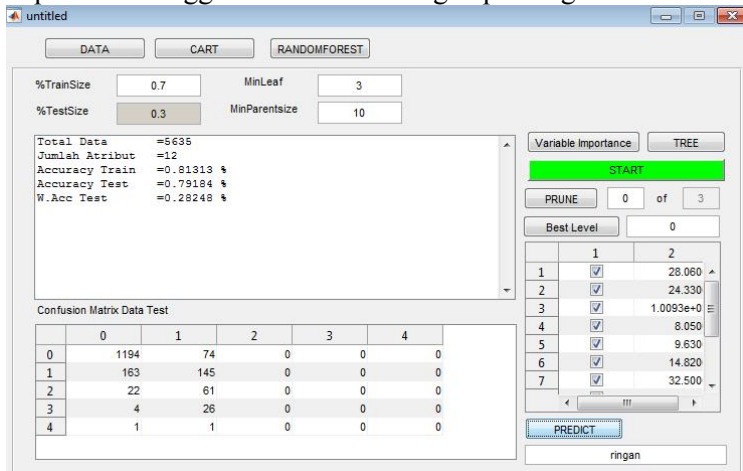
Memasukkan nilai %Trainsize yang berkisar 0-1 kemudian memasukkan nilai parameter untuk minleafsize dan minparentsiz untuk membuat model CART dengan fitctree. Menghasilkan nilai uji performa berupa confusion matrix, nilai akurasi data training, nilai akurasi data testing dan nilai weighted accuracy data testing. Setelah model dihasilkan, dapat melakukan prediksi atas kelas intensitas hujan data dengan index paling akhir pada file excel.



Gambar 5-16 Tampilan fungsi prediksi dengan metode CART

5.4.3 Fungsi Memrediksi Data dengan Metode CART Pruned

Pada Gambar 5-17 dapat dijelaskan pembuatan model pohon keputusan menggunakan CART dengan pemangkasan.

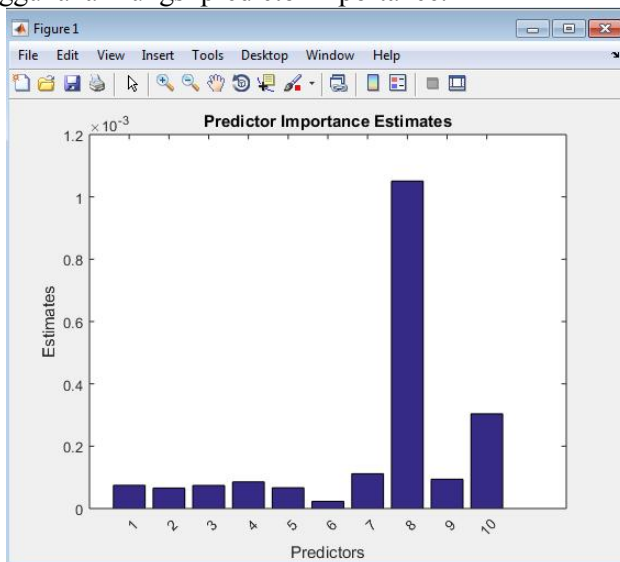


Gambar 5-17 Tampilan fungsi prediksi dengan CART pruned

Level pemangkasan dimasukkan pada kotak di sebelah tombol PRUNE, dimana dapat dihasilkan dari eksekusi tombol bestlevel yang akan menghitung level pemangkasan terbaik atau dapat kita masukkan sendiri. Menghasilkan nilai uji performa berupa confusion matrix, nilai akurasi data training, nilai akurasi data testing dan nilai weighted accuracy data testing. Setelah model dihasilkan, dapat melakukan prediksi atas kelas intensitas hujan data dengan index paling akhir pada file excel.

5.4.4 Fungsi Melihat Variable Importance CART

Pada Gambar 5-18 dapat dijelaskan tampilan antarmuka untuk menghasilkan grafik kepentingan variabel dengan menggunakan fungsi predictorImportance.



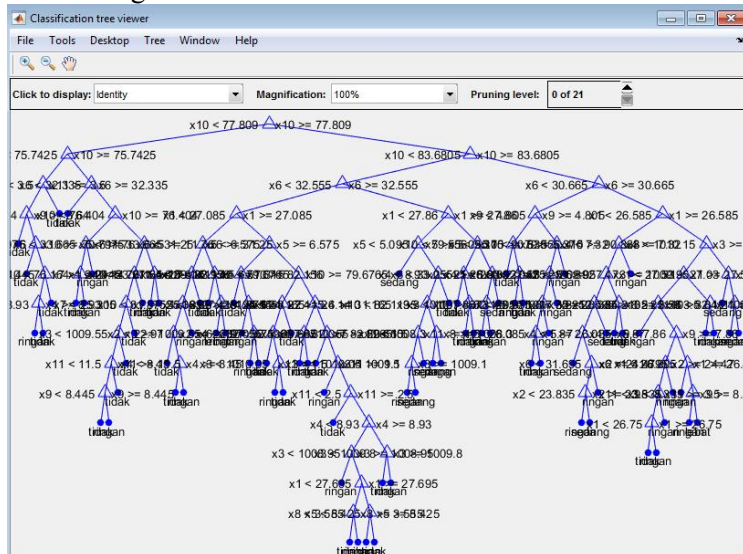
Gambar 5-18 Tampilan fungsi variabel importance CART

Hal tersebut dilakukan dengan mengeksekusi tombol Variable Importance, di mana tiap variabel akan direpresentasikan

dalam diagram batang dengan label nama variabel yang diambil dari variabel abs.

5.4.5 Fungsi Melihat Grafik Tree untuk CART

Pada Gambar 5-19 dapat dijelaskan antarmuka untuk menghasilkan grafik tree berdasarkan model CART yang telah dibuat dengan metode fitctree.

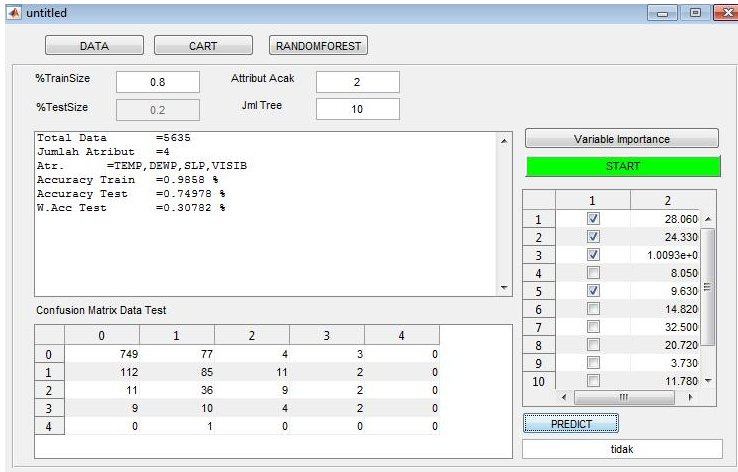


Gambar 5-19 Tampilan fungsi grafik tree untuk CART

Grafik yang dihasilkan memperlihatkan level kedalaman pohon, rule di tiap node serta label kelas pada tiap leaf node. Tiap node (kecuali leafnode) memiliki aturan keputusan terkait nilai variabel lebih dari sama dengan atau kurang dari sebuah *threshold*. Kriteria keputusan ini digunakan untuk menentukan label kelas sebuah leaf node.

5.4.6 Fungsi Memrediksi Data dengan Metode Random Forest

Pada Gambar 5-20 merupakan antarmuka pembuatan model pohon keputusan menggunakan Random Forest dengan atribut yang telah dipilih pada load data.



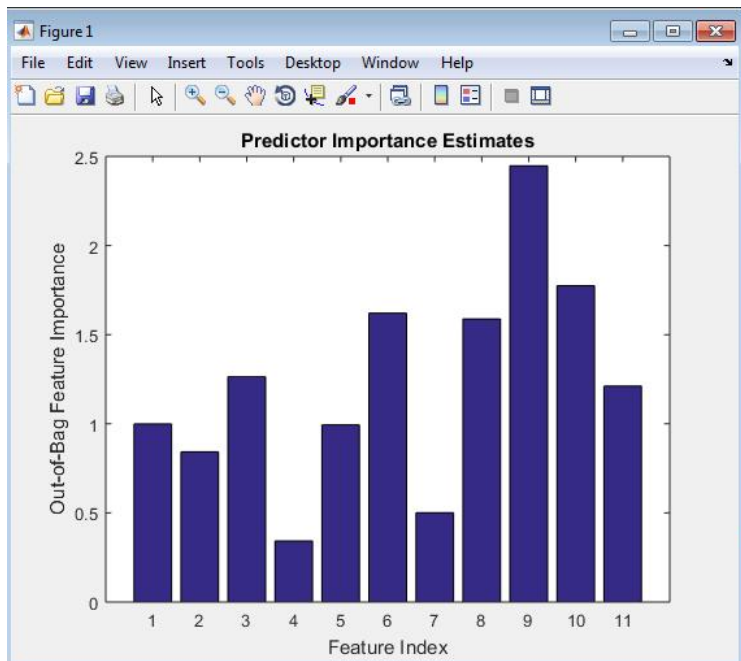
Gambar 5-20 Tampilan fungsi prediksi data dengan metode random forest

Memasukkan nilai %Trainsize yang berkisar 0-1 kemudian memasukkan nilai parameter untuk jumlah atribut acak dan jumlah pohon untuk membuat model Random Forest dengan TreeBagger. Menghasilkan nilai uji performa berupa confusion matrix, nilai akurasi data training, nilai akurasi data testing dan nilai weighted accuracy data testing. Setelah model dihasilkan, dapat melakukan prediksi atas kelas intensitas hujan data dengan index paling akhir pada file excel.

5.4.7 Fungsi Melihat Variable Importance Random Forest

Pada Gambar 5-21 adalah antarmuka untuk menghasilkan grafik kepentingan variabel dengan menggunakan fungsi predictorImportance.

Hal tersebut dilakukan dengan mengeksekusi tombol Variable Importance, di mana tiap variabel akan direpresentasikan dalam diagram batang dengan label nama variabel yang diambil dari variabel abs.



Gambar 5-21 Tampilan fungsi melihat variable importance dengan random forest

Halaman ini sengaja dikosongkan

BAB VI

HASIL DAN PEMBAHASAN

Bab ini berisikan hasil dan pembahasan setelah melakukan implementasi. Hasil yang akan dijelaskan adalah hasil uji coba model dan pembahasan tentang hal yang menyebabkan hasil yang ada terjadi.

6.1 CART

Pada tahap ini akan dilakukan analisa hasil pembuatan model dengan metode CART.

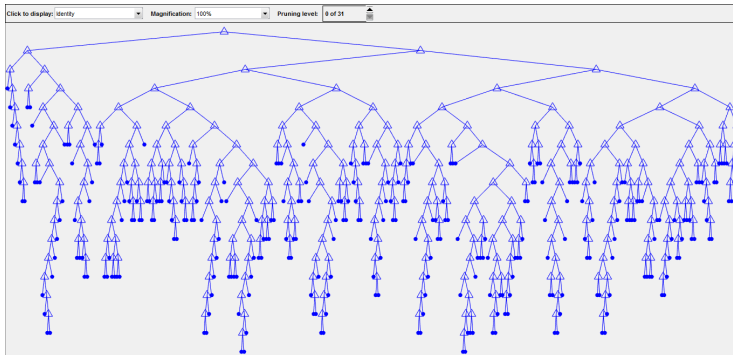
6.1.1 Model CART Maksimal

Pada tahap ini akan dilakukan analisa hasil pembuatan model dengan metode CART maksimal, dengan variasi jenis variabel.

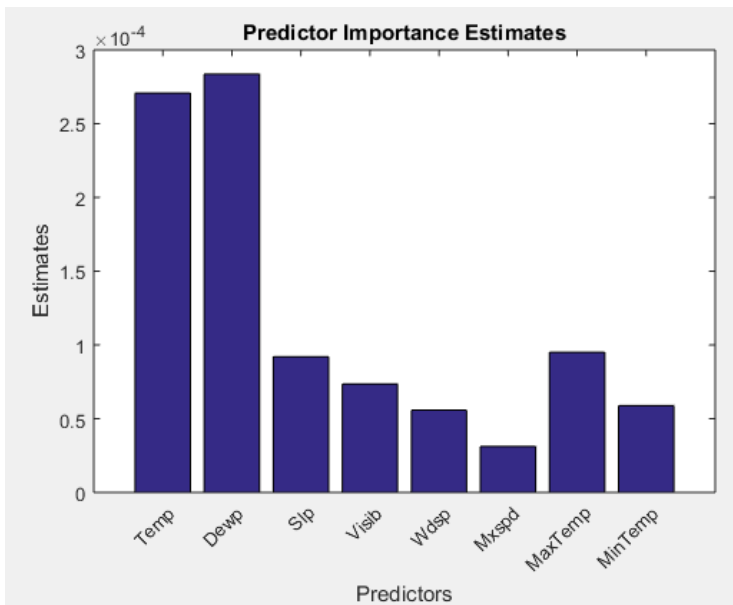
6.1.1.1 Model I - dengan 8 Variabel Awal

Pembuatan model I, menggunakan 8 variabel awal dari dataset, antara lain suhu udara, titik embun, tekanan permukaan laut, visibilitas, kecepatan angin, suhu udara minimum dan suhu udara maksimum. Decision tree yang dihasilkan memiliki level kedalaman 31, seperti yang dapat dilihat pada Gambar 6-1.

Dari 8 variabel yang digunakan untuk menghasilkan model, 2 atribut memiliki nilai kepentingan paling tinggi (di atas rata-rata) yaitu variabel suhu udara dan titik embun. Hal ini mengindikasikan nilai suhu udara dan titik embun berkontribusi besar terhadap kemampuan klasifikasi model. Suhu udara memiliki nilai kepentingan sebesar 0,2707 dan titik embun memiliki nilai kepentingan sebesar 0,2837 seperti pada Gambar 6-2.



Gambar 6-1 Decision tree Model dengan 8 variabel



Gambar 6-2 Grafik kepentingan variable untuk metode CART dengan 8 variabel

Pada Tabel 6-2 dapat dilihat nilai ukuran performa model ini untuk akurasi data training sebesar 0,9024, nilai akurasi testing bernilai sebesar 0,7291 dan nilai weghted accuracy untuk data testing sebesar 0,3015. Nilai weighted accuracy

memiliki nilai kecil karena adanya ketidakseimbangan jumlah data untuk tiap kelas, seperti dapat dilihat pada Tabel 6-1, kelas 3 dan 4 pada data test hanya memiliki 29 dan 2 data, sementara yang berhasil diprediksi secara benar untuk kelas 3 hanya 2/29 dan untuk kelas 4 tidak ada yang berhasil diprediksi.

Tabel 6-1 Confusion Matrix Model I dengan 8 variabel

	0	1	2	3	4	Actual Class
0	1084	133	25	7	0	1249
1	138	134	26	12	0	310
2	26	54	14	7	0	101
3	8	12	7	2	0	29
4	1	0	1	0	0	2

Tabel 6-2 Nilai akurasi CART Model I (8 variabel)

Ukuran Performa	Nilai
Akurasi Training	0.9204
Akurasi Testing	0.7297
Weighted Accuracy Testing	0.3015

Selain dari segi confusion matrix, dari segi nilai ukuran performa lain seperti precision dan recall. Seperti dapat dilihat pada Tabel 6-3, yang menjadi perhatian khusus adalah kelas 3 dan kelas 4. Precision kelas 3 hanya 0,071 sementara kelas 4 tidak dapat dihitung yang menunjukkan ketepatan pengklasifikasi yang cukup buruk. Recall kelas 3 bernilai 0,069 dan kelas 4 bernilai 0,00 yang menunjukkan kelengkapan pengklasifikasi yang cukup buruk. Pada model ini kelas 4 belum dapat diprediksi oleh model.

Tabel 6-3 Nilai precision dan recall CART Model I (8 variabel)

Kelas	Recall	Precision
0	0,868	0,862
1	0,432	0,162

Kelas	Recall	Precision
2	0,139	0,192
3	0,069	0,071
4	0,000	#DIV/0!

6.1.1.2 Model dengan Varibel hasil Pemilihan Variable Importance dari Model I

Dari 2 variabel yang digunakan untuk menghasilkan model selanjutnya, dimana hanya menggunakan variabel yang memiliki nilai kepentingan di atas rata-rata seperti dapat dilihat pada Tabel 6-4, yaitu suhu udara dan titik embun. Nilai kepentingan yang dipilih adalah yang lebih besar dari nilai rata-rata, yaitu sebesar 0,1201.

Tabel 6-4 Kepentingan variabel dari Model I (8 variabel)

No	Variable	Variable Importance
	Temp	0,2707
	Dewp	0,2837
	Slp	0,0920
	Visib	0,0735
	Wdsp	0,0559
	Mxspd	0,0312
	MaxTemp	0,0950
	MinTemp	0,0589

Mean = 0,1201

Seperti dapat dilihat pada Tabel 6-5 nilai ukuran performa model ini untuk akurasi data training sebesar 0,8596, nilai akurasi testing bernilai sebesar 0,7363 dan nilai weighted accuracy untuk data testing sebesar 0,2856. Nilai weighted accuracy memiliki nilai kecil karena adanya ketidakseimbangan jumlah data untuk tiap kelas. Semakin banak proporsi data yang tidak dapat diprediksi dengan benar, maka semakin kecil nilai weighted accuracy.

Tabel 6-5 Nilai akurasi CART Model I (dengan pemilihan atribut)

Ukuran Performa	Nilai
Akurasi Training	0.8596
Akurasi Testing	0.7363
Weighted Accuracy Testing	0.2856

Selain dari segi confusion matrix, dari segi nilai ukuran performa lain seperti precision dan recall. Seperti dapat dilihat pada Tabel 6-6, yang menjadi perhatian khusus adalah kelas 3 dan kelas 4. Precision kelas 3 sebesar 0,00 sementara kelas 4 tidak dapat dihitung yang menunjukkan ketepatan pengklasifikasi yang buruk. Recall kelas 3 bernilai 0,000 dan kelas 4 bernilai 0,00 yang menunjukkan kelengkapan pengklasifikasi yang buruk. Pada model ini baik kelas 3 maupun kelas 4 belum dapat diprediksi oleh model.

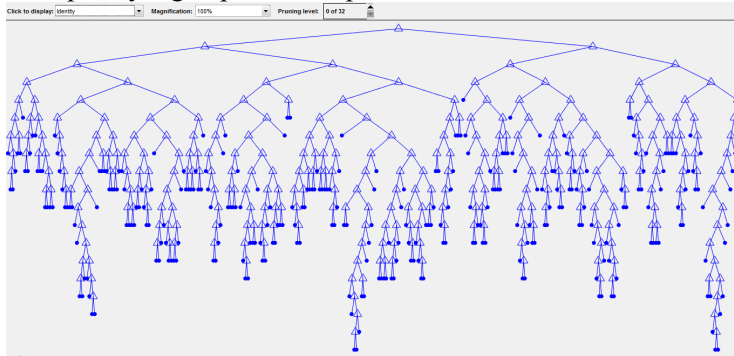
Tabel 6-6 Nilai precision dan recall CART Model I (dengan pemilihan atribut)

Kelas	Recall	Precision
0	0,883	0,856
1	0,416	0,175
2	0,129	0,250
3	0,000	0,000
4	0,000	#DIV/0!

6.1.1.3 Model II - dengan 8 Variabel Awal dan 4 Variabel Turunan

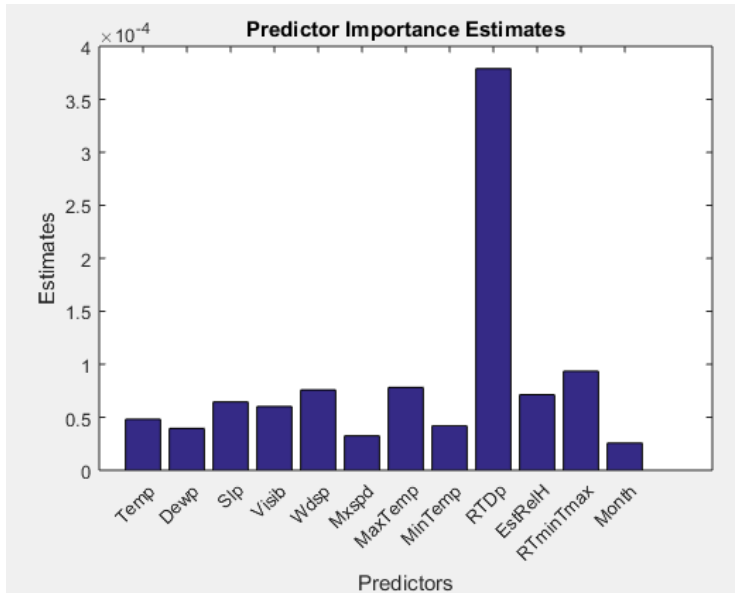
Pembuatan model I, menggunakan 8 variabel awal dari dataset, antara lain suhu udara, titik embun, tekanan permukaan laut, visibilitas, kecepatan angin, suhu udara minimum dan suhu udara maksimum serta 4 variabel turunan yaitu selisih suhu udara dan titik embun, estimasi kelembapan

relatif, selisih suhu udara maksimum dan minimum serta bulan. Decision tree yang dihasilkan memiliki level kedalaman 32 seperti yang dapat dilihat pada Gambar 6-3.



Gambar 6-3 Grafik tree untuk model CART dengan 12 variabel

Dari 12 variabel yang digunakan untuk menghasilkan model, 2 atribut memiliki nilai kepentingan paling tinggi (di atas rata-rata) yaitu variabel selisih suhu udara dan titik embun serta selisih suhu udara maksimum dan minimum. Hal ini mengindikasikan nilai selisih suhu udara dan titik embun serta selisih suhu udara maksimum dan minimum berkontribusi besar terhadap kemampuan klasifikasi model. selisih suhu udara dan titik embun memiliki nilai kepentingan sebesar 0,3789 serta selisih suhu udara maksimum dan minimum memiliki nilai kepentingan sebesar 0,0936. Nilai kepentingan ini bersifat relatif, karena akan berubah bila dibandingkan dengan variabel yang lainnya (dikurangi maupun ditambahi variabelnya). Nilai selisih suhu udara dan titik embun memiliki nilai yang paling menonjol dibandingkan variabel lain, seperti yang dapat dilihat pada Gambar 6-4.



Gambar 6-4 Grafik kepentingan variabel untuk metode CART dengan 12 variabel

Seperti dapat dilihat pada Tabel 6-7, nilai ukuran performa model ini untuk akurasi data training sebesar 0,9321, nilai akurasi testing bernilai sebesar 0,7286 dan nilai weighted accuracy untuk data testing sebesar 0,3216. Nilai weighted accuracy memiliki nilai kecil karena adanya ketidakseimbangan jumlah data untuk tiap kelas. Semakin banyak proporsi data yang tidak dapat diprediksi dengan benar, maka semakin kecil nilai weighted accuracy.

Tabel 6-7 Nilai akurasi CART Model I (8 variabel awal dan 4 variabel baru)

Ukuran Performa	Nilai
Akurasi Training	0.9321
Akurasi Testing	0.7286
Weighted Accuracy Testing	0.3216

Selain dari segi confusion matrix, dari segi nilai ukuran performa lain seperti precision dan recall. Seperti dapat dilihat pada Tabel 6-8, yang menjadi perhatian khusus adalah kelas 3 dan kelas 4. Precision kelas 3 hanya 0,192 sementara kelas 4 tidak dapat dihitung yang menunjukkan ketepatan pengklasifikasi yang cukup buruk. Recall kelas 3 bernilai 0,172 dan kelas 4 bernilai 0,00 yang menunjukkan kelengkapan pengklasifikasi yang cukup buruk. Pada model ini kelas 4 belum dapat diprediksi oleh model.

Tabel 6-8 Nilai precision dan recall CART (8 variabel awal dan 4 variabel baru)

Kelas	Recall	Precision
0	0,861	0,870
1	0,445	0,176
2	0,129	0,186
3	0,172	0,192
4	0,000	#DIV/0!

6.1.1.4 Model dengan Variabel hasil Pemilihan Variable Importance dari Model II

Dari 2 variabel yang digunakan untuk menghasilkan model selanjutnya, dimana hanya menggunakan variabel yang memiliki nilai kepentingan di atas rata-rata seperti dapat dilihat pada Tabel 6-9, yaitu RTDp (selisih suhu udara dan titik embun) dan R_{tmin}T_{max} (selisih suhu udara maksimum dan minimum). Nilai kepentingan yang dipilih adalah yang lebih besar dari nilai rata-rata, yaitu sebesar 0,0842.

Tabel 6-9 Kepentingan variabel Model II (8 variabel awal dan 4 variabel baru)

No	Variable	Variable Importance
1.	Temp	0,0482
2.	Dewp	0,0394

No	Variable	Variable Importance
3.	Slp	0,0645
4.	Visib	0,0601
5.	Wdsp	0,0759
6.	Mxspd	0,0326
7.	MaxTemp	0,0782
8.	MinTemp	0,0419
9.	RTDp	0,3789
10.	EstRelH	0,0714
11.	RtminTmax	0,0936
12.	Month	0,0257

Mean = 0,0842

Pada Tabel 6-10, nilai ukuran performa model ini untuk akurasi data training sebesar 0,8781, nilai akurasi testing bernilai sebesar 0,7280 dan nilai weighted accuracy untuk data testing sebesar 0,3146. Nilai weighted accuracy memiliki nilai kecil karena adanya ketidakseimbangan jumlah data untuk tiap kelas. Semakin banyak proporsi data yang tidak dapat diprediksi dengan benar, maka semakin kecil nilai weighted accuracy.

Tabel 6-10 Nilai akurasi CART Model II (8 variabel awal dan 4 variabel baru)

Ukuran Performa	Nilai
Akurasi Training	0.8781
Akurasi Testing	0.7280
Weighted Accuracy Testing	0.3146

Selain dari segi confusion matrix, dari segi nilai ukuran performa lain seperti precision dan recall. Seperti dapat dilihat pada Tabel 6-11, yang menjadi perhatian khusus adalah kelas 3 dan kelas 4. Precision kelas 3 sebesar 0,167 sementara kelas 4 tidak dapat dihitung yang menunjukkan ketepatan pengklasifikasi yang cukup buruk. Recall kelas 3 bernilai 0,172 dan kelas 4 bernilai 0,00 yang menunjukkan kelengkapan pengklasifikasi yang cukup buruk. Pada model ini baik kelas 4 belum dapat diprediksi oleh model.

Tabel 6-11 Nilai precision dan recall Model II (8 variabel awal dan 4 variabel baru)

Kelas	Recall	Precision
0	0,863	0,860
1	0,448	0,160
2	0,089	0,176
3	0,172	0,167
4	0,000	#DIV/0!

6.1.1.5 Perbandingan Model CART Maksimal

Seperti dapat dilihat pada Tabel 6-12, model I (dengan pemilihan atribut) memiliki nilai ukuran performa akurasi data testing yang paling tinggi yaitu 0,7363.

Tabel 6-12 Tabel perbandingan nilai akurasi Model CART maksimal

Ukuran Performa	Model I (8 Variabel)	Model I dengan pemilihan Atribut	Model II (12 Variabel)	Model II dengan pemilihan atribut
Akurasi Training	0.9204	0.8596	0.9321	0.8781
Akurasi Testing	0.7297	0.7363	0.7286	0.7280
Weighted Accuracy Testing	0.3015	0.2856	0.3216	0.3146

Namun memiliki nilai precision dan recall untuk yang buruk karena tidak mampu memprediksi satu pun kelas 3 dan kelas 4 yang benar, sementara model lain mampu memprediksi hingga kelas 3 (kelas 4 tidak mampu diprediksi oleh semua model).

Tabel 6-13 Nilai precision dan recall Model CART maksimal

K E L A S	Model I (8 Variabel)		Model I dengan pemilihan Attribut		Model II (12 Variabel)		Model II dengan pemilihan Attribut	
	Recall	Preci sion	Rec all	Preci sion	Rec all	Preci sion	Rec all	Preci sion
0	0,868	0,862	0,88 3	0,856	0,86 1	0,870	0,86 3	0,860
1	0,432	0,162	0,41 6	0,175	0,44 5	0,176	0,44 8	0,160
2	0,139	0,192	0,12 9	0,250	0,12 9	0,186	0,08 9	0,176
3	0,069	0,071	0,00 0	0,000	0,17 2	0,192	0,17 2	0,167
4	0,000	#DIV/ 0!	0,00 0	#DIV/ 0!	0,00 0	#DIV/ 0!	0,00 0	#DIV/ 0!

Model yang terbaik dipilih adalah model dengan jumlah variabel 12, karena memiliki nilai relatif lebih besar dibandingkan dengan model yang lain, untuk nilai akurasi training paling tinggi yaitu 0.9321, untuk nilai akurasi testing paling tinggi yaitu 0.7286, nilai weighted accuracy paling tinggi yaitu 0.3216 serta mampu memprediksi kelas 0,1,2 dan 3 jika dilihat dari precision dan recall seperti dapat dilihat pada Tabel 6-13.

Berdasarkan variabel yang terpilih berdasarkan nilai variable importance, variabel yang paling berpengaruh antara lain suhu udara, titik embun, selisih suhu udara dan titik embun serta selisih suhu udara maksimum dan minimum yang menunjukkan 2 dari 4 variabel turunan yaitu selisih suhu udara dan titik embun dan selisih suhu udara maksimum dan minimum, memiliki nilai signifikan terhadap peningkatan nilai performa model.

6.1.2 Model dengan Pruning hasil Model CART Maksimal

Setelah dipilih variabel untuk menghasilkan model paling baik, kemudian melakukan pemangkasan decision tree yang telah dihasilkan. Pada model dengan 12 variabel dengan minleaf dan minparent default, yaitu :

1. Minleafsize (default: 10)
2. MinParentSize (default: 1)

Dengan proporsi data training 70% dan data testing 30%

bestlevel = 29

Mencari nilai level pemangkasan terbaik berdasarkan nilai complexity paling minimal. Pembuatan pohon yang optimal dengan melakukan pemangkasan dengan mencari nilai penduga pengganti validasi silang sebanyak 3 fold (3 bagian). Pemilihan umlah fold sebanyak 3 dengan mempertimbangkan jumlah data pada kelas yang paling sedikit memiliki data, yaitu kelas 4 yang memiliki jumlah sebanyak 5 data pada data training, sehingga tiap fold dapat memiliki bagan data dari kelas data terkecil.

Tabel 6-14 Nilai akurasi CART Pruned

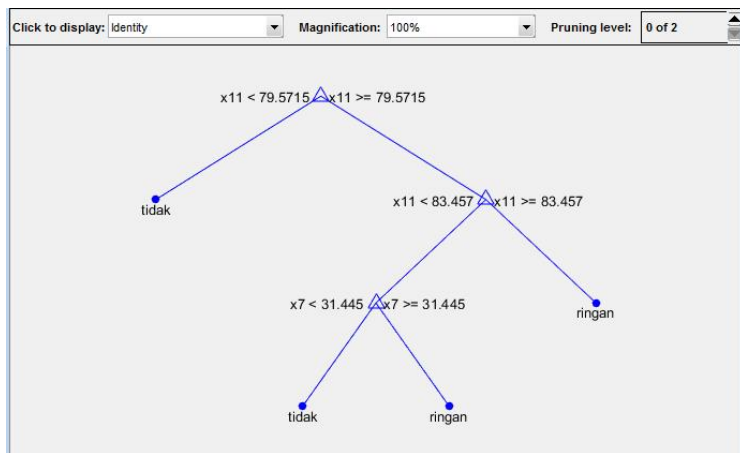
Ukuran Performa	Nilai
Akurasi Training	0.8067
Akurasi Testing	0.7877
Weighted Accuracy Testing	0.2937

Seperti dapat dilihat pada Tabel 6-14, nilai akurasi data testing untuk pohon keputusan yang telah dipangkas naik 0,6 dibandingkan dengan model CART maksimal dengan 12 variabel (yang terbaik). Namun, peningkatan akurasi ini justru memiliki tradeoff terhadap kemampuan model untuk memprediksi multi kelas. Dari hasil pengukuran precision seperti dapat dilihat pada

Tabel 6-15 dan recall, diketahui bahwa model CART yang dipangkas hanya bisa memprediksi kelas 0 dan 1.

Tabel 6-15 Nilai precision dan recall CART Pruned

Kelas	Recall	Precision
0	0,916	0,879
1	0,552	0,179
2	0,000	#DIV/0!
3	0,000	#DIV/0!
4	0,000	#DIV/0!



Gambar 6-5 Pohon keputusan hasil CART dengan pemangkas

Dari Gambar 6-5 menunjukkan hasil pohon keputusan setelah dipangkas menjadi 2 level, dapat diketahui beberapa aturan untuk klasifikasi kelas hujan, antara lain: x_{11} -RELH (estimasi kelembapan relatif) , x_7 (Suhu udara maksimum)

1. Jika RELH kurang dari 79.5715 maka tidak akan terjadi hujan

2. Jika RELH lebih dari sama dengan 83.457 maka akan terjadi hujan ringan
3. Jika RELH berada di antara 79.5715 dan 83.457, serta Suhu udara maksimum kurang dari 31.445, maka tidak akan terjadi hujan
4. Jika RELH berada di antara 79.5715 dan 83.457, serta Suhu udara maksimum lebih dari sama dengan 31.445, maka akan terjadi hujan ringan

6.1.3 Model dengan Pengendalian Kedalaman Pohon

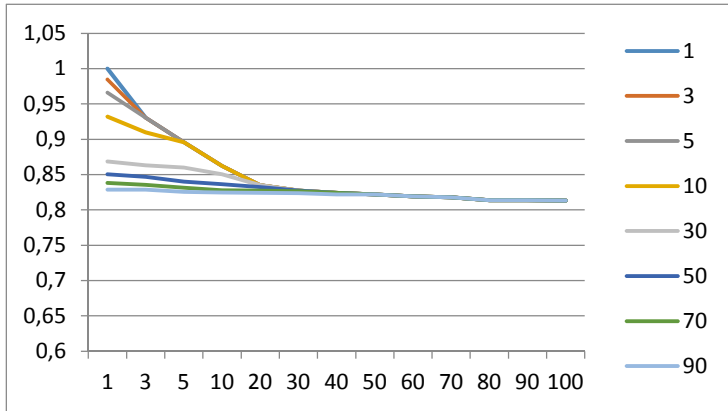
Pada tahap ini akan dilakukan analisa hasil pembuatan model dengan metode CART dengan pengendalian kedalaman pohon, yaitu melakukan tuning parameter jumlah minimal leaf node dan jumlah minimal parent node.

6.1.3.1 Hasil Uji Performa Berdasarkan Akurasi Training, Akurasi Testing dan Weighted Accuracy

Pada tahap ini dianalisa hasil uji performa model dengan kombinasi jumlah minimal leaf node dan jumlah minimal parent node dari segi nilai akurasi training, akurasi testing dan nilai weighted accuracy.

A. Akurasi Training

Pada bagian ini dijelaskan nilai performa akurasi training.

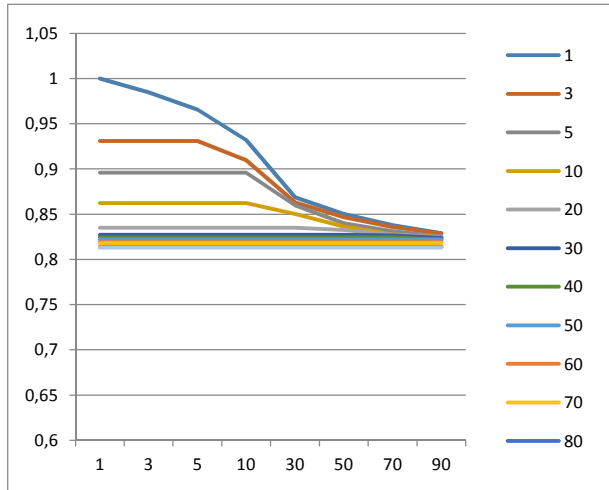


Gambar 6-6 Grafik perbandingan jumlah Minleaf (x) terhadap Minparentsize (y) CART untuk akurasi training

Pada Gambar 6-6 menunjukkan akurasi data training paling tinggi dengan jumlah minleaf 1 dan semakin menurun hingga mencapai kestabilan pada leaf 20 ke 100 yaitu pada nilai akurasi sebesar 0,80 hingga 0,85.

Pada Gambar 6-7 menunjukkan akurasi data training paling tinggi dengan jumlah parent node 1 dan semakin menurun hingga mencapai kestabilan pada parent node 50 ke atas yaitu pada nilai akurasi sebesar 0,80 hingga 0,85.

Sehingga dengan nilai leaf 20 hingga 100, dan parent size 50 hingga 90 mencapai kestabilan akurasi training sebesar 0,8 hingga 0,85. Nilai training lebih detail dapat dilihat pada Tabel C-1.

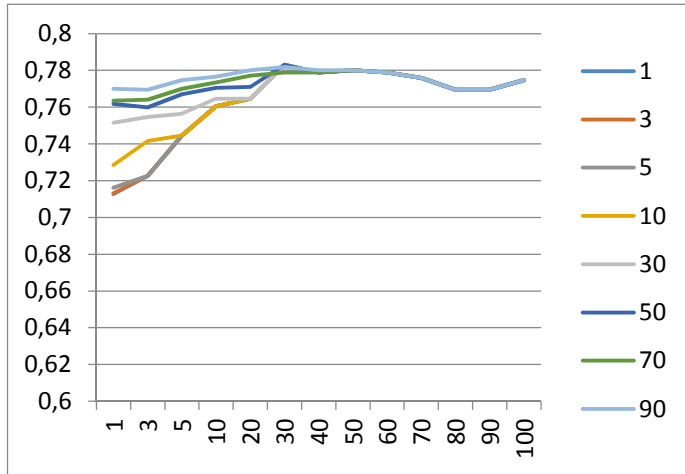


Gambar 6-7 Grafik perbandingan jumlah Minparentsize (x) terhadap Minleaf (y) CART untuk akurasi training

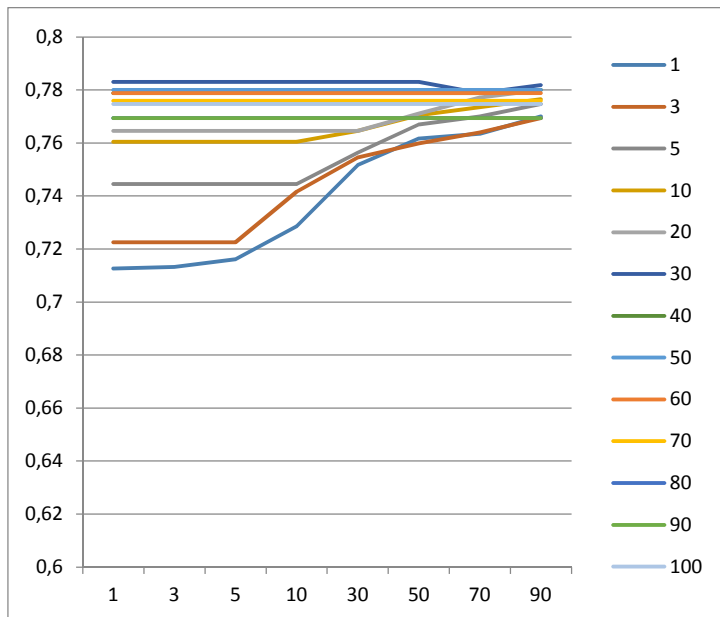
B. Akurasi Testing

Pada bagian ini dijelaskan nilai performa akurasi testing.

Pada Gambar 6-8 menunjukkan akurasi data testing paling tinggi dengan jumlah minleaf 1 dan semakin naik hingga mencapai kestabilan pada leaf 30 hingga 60 yaitu pada nilai akurasi sebesar 0,78.



Gambar 6-8 Grafik perbandingan jumlah Minleaf (x) terhadap Minparentsiz (y) CART untuk akurasi testing



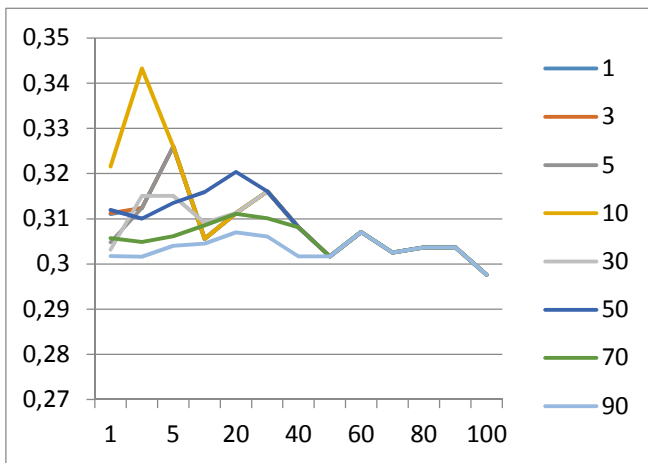
Gambar 6-9 Grafik perbandingan jumlah Minparentsize (x) terhadap Minleaf (y) CART untuk akurasi testing

Pada Gambar 6-9 menunjukkan model mencapai kestabilan pada parent node 50 ke 90 yaitu pada nilai akurasi sebesar 0,76 hingga 0,78.

Sehingga dengan nilai leaf 30 hingga 60, dan parent size 50 hingga 90 mencapai kestabilan akurasi testing sebesar 0,78. Nilai training lebih detail dapat dilihat pada Tabel C-2.

C. Weighted Accuracy Testing

Pada bagian ini dijelaskan nilai performa weighted akurasi testing.



Gambar 6-10 Grafik perbandingan jumlah Minleaf (x) terhadap Minparentsize (y) CART untuk Weighted Accuracy Testing

Pada Gambar 6-10 menunjukkan akurasi data paling tinggi dengan jumlah minleaf 1 dan semakin turun hingga mencapai kestabilan pada leaf 50 hingga 90 yaitu pada nilai akurasi

sebesar 0,3. Nilai weighted accuracy lebih detail dapat dilihat pada Tabel C-3.

6.1.3.2 Hasil Uji Performa Berdasarkan Presisi dan Recall

Pada tahap ini dianalisa hasil uji performa model dengan kombinasi jumlah minimal leaf node dan jumlah minimal parent node dari segi nilai nilai presisi dan recall.

A. Presisi Model untuk Kelas 0,1,2

Pada kelas 0,1,2 semua model uji dapat mengklasifikasi dengan baik seperti terlihat pada Tabel C-4, Tabel C-5 dan Tabel C-6, atau paling tidak terdapat kelas yang diprediksi dengan benar.

B. Presisi Model untuk Kelas 3

Seperti dapat dilihat pada Tabel 6-16, Model yang mampu memprediksi kelas 3 dengan benar hanya model dengan leaf dari 1 sampai 10 serta parent size dari 1 sampai 50. Selain model tersebut, tidak dapat sama sekali memprediksi kelas 3, yang ditunjukkan dengan NaN yaitu nilai yang dihasilkan karena nilai prediksi benar sama dengan 0 yang dibagi dengan total kelas prediksi yang juga bernilai 0.

Tabel 6-16 Nilai presisi CART (Pengendalian kedalaman pohon) untuk kelas 3

Minimal Leaf	Minimal Parent								
		1	3	5	10	30	50	..	90
	1	0,179	0,156	0,125	0,192	0,111	0,143	..	NaN
	3	0,160	0,160	0,160	0,240	0,182	0,143	..	NaN
	5	0,231	0,231	0,231	0,231	0,200	0,167	..	NaN
	10	0,167	0,167	0,167	0,167	0,167	0,167	..	NaN
	:							..	NaN
	100	NaN	NaN	NaN	NaN	NaN	NaN	..	NaN

Nilai leaf 3 dan 5 memiliki nilai presisi yang cenderung lebih tinggi dibandingkan leaf yang lainnya.

C. Presisi Model untuk Kelas 4

Tidak terdapat model yang mampu memprediksi kelas 4 dengan benar hanya model dengan leaf 1 serta parent size dari 1 sampai 5 mampu memprediksi kelas 4 meskipun tidak ada hasil prediksi yang benar, yang ditunjukkan nilai 0, yaitu hasil kelas prediksi (0) dibagi total kelas prediksi (>1). Selain model tersebut, tidak dapat sama sekali memprediksi kelas 4, yang ditunjukkan dengan NaN yaitu nilai yang dihasilkan karena nilai prediksi benar sama dengan 0 yang dibagi dengan total kelas prediksi. Nilai presisi model ntuk kelas 4 lebih detail dapat dilihat pada Tabel C-7.

D. Recall Model untuk Kelas 0,1,2

Pada kelas 0,1,2 semua model uji dapat mengklasifikasi dengan baik, atau paling tidak terdapat kelas yang diprediksi dengan benar. Nilai recall kelas 0,1,2 lebih detail dapat dilihat pada Tabel C-8, Tabel C-9, Tabel C-10.

E. Recall Model untuk Kelas 3

Seperti dapat dilihat pada Tabel 6-17, model yang mampu memprediksi kelas 3 dengan benar hanya model dengan leaf dari 1 sampai 10 serta parent size dari 1 sampai 50. Selain model tersebut, tidak dapat sama sekali memprediksi kelas 3, yang ditunjukkan dengan 0 yaitu nilai yang dihasilkan karena nilai prediksi benar sama dengan 0 yang dibagi dengan total kelas prediksi yang bernilai ≥ 1 .

Tabel 6-17 Nilai recall CART (Pengendalian kedalaman pohon) untuk kelas 3

Minimal Leaf	Minimal Parent							
		1	3	5	10	30	50	... 90
	1	0,172	0,172	0,138	0,172	0,035	0,0345	... 0
	3	0,138	0,138	0,138	0,207	0,069	0,0345	... 0
	5	0,103	0,103	0,103	0,103	0,069	0,0345	... 0
	10	0,035	0,035	0,035	0,035	0,034	0,0345	... 0
	:	:	:	:	:	:	:	... 0
	100	0	0	0	0	0	0	... 0

Nilai leaf 1 dan 3 memiliki nilai presisi yang cenderung lebih tinggi dibandingkan leaf yang lainnya.

F. Recall Model untuk Kelas 4

Tidak terdapat model yang mampu memprediksi kelas 4 dengan benar, yang ditunjukkan nilai 0, yaitu hasil kelas prediksi (0) dibagi total kelas prediksi (>1). Selain model tersebut, tidak dapat sama sekali memprediksi kelas 4, yang ditunjukkan dengan NaN yaitu nilai yang dihasilkan karena nilai prediksi benar sama dengan 0 yang dibagi dengan total kelas prediksi. Nilai recall model kelas 4 lebih detail dapat dilihat pada Tabel C-11.

6.1.4 Pengujian Model CART dengan Proporsi Presentase Data Training dan Testing

Pada Tabel 6-18 menunjukkan untuk semua kombinasi proporsi data training dan testing, menghasilkan akurasi yang cenderung stabil, yaitu untuk data training berkisar antara 0,85-0,86 dan untuk data testing berkisar antara 0,76-0,775 dengan 0,775 dimiliki oleh proporsi data training-testing sebesar 60%-40%. Pada nilai Weighted Accuracy, cenderung stabil berkisar antara 0,30-0,32.

Tabel 6-18 Nilai akurasi pengujian proporsi CART

Training	Testing	Tr	Test	WaccT
60%	40%	0,858	0,775	0,319
70%	30%	0,857	0,767	0,324
75%	25%	0,855	0,768	0,323
80%	20%	0,861	0,748	0,323
85%	15%	0,862	0,763	0,320
90%	10%	0,858	0,761	0,300

6.1.5 Kesimpulan Hasil Model CART

Berikut ini adalah kesimpulan dari model CART:

1. Variabel yang dipilih untuk dilibatkan dalam pembuatan model adalah 12 variabel (8 variabel dataser, 4 variabel turunan) karena memiliki nilai uji performa yang baik dari segi akurasi dan dari segi presisi dan recall. Model dengan 12 variabel ini menggunakan parameter default minimal leaf sejumlah 10 dan minimal parent size 1 dan proporsi training/testing 70%/30%.
2. Model dengan 12 variabel memiliki nilai akurasi training yaitu 0.9321, untuk nilai akurasi testing yaitu 0.7286, nilai weighted accuracy yaitu 0.3216 serta mampu memprediksi kelas 0,1,2,3 jika dilihat dari precision dan recall, namun menghasilkan decision tree 31 level.
3. Pemangkasan CART maksimal dengan Model 12 variabel menghasilkan 2 level decision tree, menyisakan 2 variabel yaitu x11-RELH (estimasi kelembapan relatif) , x7 (Suhu udara maksimum) dengan nilai akurasi training yaitu 0.8067, untuk nilai akurasi testing yaitu 0.7877, nilai weighted accuracy yaitu 0.2937 serta hanya mampu memprediksi kelas 0,1

4. Pengaturan kedalaman pohon, pada Tabel 6-19 adalah hasil sesuai hasil masing-masing uji performa, dibagi menjadi 2 bagian yaitu hasil yang stabil dan yang tinggi:

Tabel 6-19 Nilai parameter berdasarkan uji performa pengendalian kedalaman pohon model CART

	Tra in	Test	Wacc Test	Presisi (memprediks i 0,1,2,3)	Recall (memprediks i 0,1,2,3)
<i>Stabil</i>					
<i>Leaf</i>	20-100	30-60	50-90		
<i>Parent</i>	50-90	50-90	All		
<i>Tinggi</i>					
<i>Leaf</i>	1-5	30-60	1-5	3-5	1-5
<i>Parent</i>	1-10	50-90	10	1-50	1-50

Sehingga, model yang paling baik yang dapat memprediksi 0,1,2,3 adalah dengan leaf (1-5) dan parent (30). Model tersebut pada sisi presisi (ketepatan pengklasifikasi) untuk kelas 0 memiliki perfroma 0.868, kelas 1 sebesar 0.422, kelas 2 sebesar 0.438, kelas 3 sebesar 0.182 dan kelas 4 tidak berhasil diklasifikasi. Dari segi recall (kelengkapan pengklasifikasi) untuk kelas 0 memiliki perfomas sebesar 0.899, kelas 1 sebesar 0.429, kelas 2 sebesar 0.178 , kelas 3 sebesar 0.069 dan kelas 4 tidak bisa diklasifikasikan dengan benar. Hal ini menunjukkan kemampuan pengklasifikasi model paling baik pada kelas 0, cukup baik pada kelas 1 (hanya sekitar 40% benar) , pada kelas 2 dan 3 hanya berkisar 6%-20% yang benar.

5. Proporsi training dan testing untuk model CART dengan leaf 3 dan parent 30 untuk proporsi 60:40, 70:30, 75:25, 80:20, 85:15 dan 90:10 cenderung stabil, untuk data training berkisar antara 0,85-0,86 dan untuk data testing berkisar antara 0,76-0,775 serta nilai Weighted Accuracy, cenderung stabil berkisar antara 0,30-0,32.
6. Melakukan pengujian kestabilan proporsi data training/testing untuk model CART dengan leaf 3 dan parent 30 menunjukkan proporsi 60% data training dan 40% data testing dengan nilai akurasi training 0,858, akurasi testing 0,775 dan weighted accuracy testing 0,318.

6.2 Random Forest

Pada tahap ini akan dilakukan pengolahan model dengan menggunakan teknik Random Forest.

6.2.1 Pemilihan Atribut Model

Parameter model diatur default yaitu dengan jumlah variable acak sebesar 2. Eksperimen model dilakukan dengan melihat nilai uji performa pada jenis variabel yang dilibatkan dalam pembuatan model dan jumlah tree untuk melihat kestabilan model.

6.2.1.1 Model I - dengan 8 Variabel Awal

Pembuatan model I, menggunakan 8 variabel awal dari dataset, antara lain suhu udara, titik embun, tekanan permukaan laut, visibilitas, kecepatan angin, suhu udara minimum dan suhu udara maksimum.

Seperti dapat dilihat pada

Tabel 6-20, nilai ukuran performa model ini untuk akurasi data training stabil sebesar 1, nilai akurasi testing bernilai sebesar berkisar antara 0,77-0,78 dan nilai weighted accuracy untuk data testing berkisar 0,3-0,31. Nilai weighted accuracy

memiliki nilai kecil karena adanya ketidakseimbangan jumlah data untuk tiap kelas.

Tabel 6-20 Nilai akurasi Random Forest Model I (8 variabel)

nTree	Training	Testing	Wacc Testing
50	1	0,78	0,3115
150	1	0,7765	0,3079
250	1	0,7776	0,309

Selain dari segi confusion matrix, dari segi nilai ukuran performa lain seperti precision dan recall. Seperti dapat dilihat pada Tabel 6-20, yang menjadi perhatian khusus adalah kelas 3 dan kelas 4. Precision kelas 3 tidak dapat dihitung dan kelas 4 juga tidak dapat dihitung yang menunjukkan ketepatan pengklasifikasi yang sangat buruk. Recall kelas 3 bernilai 0 dan kelas 4 bernilai 0 yang menunjukkan kelengkapan pengklasifikasi yang cukup buruk. Pada model ini kelas 3 dan 4 belum dapat diprediksi oleh model.

Tabel 6-21 Nilai precision dan recall Random forest Model I (8 variabel)

Kelas	50		150		250	
	Recall	Precision	Recall	Precision	Recall	Precision
0	0,9322	0,863404	0,939	0,870	0,935	0,866
1	0,461	0,179054	0,494	0,174	0,481	0,176
2	0,0723	0,230769	0,084	0,292	0,060	0,238
3	0	#DIV/0!	0	#DIV/0!	0	0
4	0	#DIV/0!	0	#DIV/0!	0	#DIV/0!

6.2.1.2 Model dengan Variabel hasil Pemilihan Variable Importance dari Model I

Dari 2 variabel yang digunakan untuk menghasilkan model selanjutnya, dimana hanya menggunakan variabel yang

memiliki nilai kepentingan di atas rata-rata, seperti dapat dilihat pada Tabel 6-22, yaitu suhu udara dan titik embun. Nilai kepentingan yang dipilih adalah yang lebih besar dari nilai rata-rata.

Tabel 6-22 Kepentingan variabel Random forest Model I (8 variabel)

Index	50	150	250
1	4,412	4,486	4,580
2	4,344	4,363	4,099
3	1,178	1,017	1,148
4	0,756	0,921	0,828
5	0,631	0,876	0,812
6	0,565	0,642	0,580
7	1,209	1,439	1,273
8	1,136	0,974	1,050

Mean tree 50 = 1,779

Mean tree 150 = 1,840

Mean tree 250 = 1,796

Dari 2 variabel yang digunakan untuk menghasilkan model selanjutnya, dimana hanya menggunakan variabel yang memiliki nilai kepentingan di atas rata-rata, yaitu suhu udara dan titik embun. Nilai kepentingan yang dipilih adalah yang lebih besar dari nilai rata-rata, yaitu berkisar antara 1,77-1,8.

Seperti dapat dilihat pada Tabel 6-23, nilai ukuran performa model ini untuk akurasi data training berkisar antara 0,93 , nilai akurasi testing bernilai berkisar pada 0,72 dan nilai weighted accuracy untuk data testing berkisar antara 0,28-0,29. Nilai weighted accuracy memiliki nilai kecil karena adanya ketidakseimbangan jumlah data untuk tiap kelas.

Semakin banyak proporsi data yang tidak dapat diprediksi dengan benar, maka semakin kecil nilai weighted accuracy.

Tabel 6-23 Nilai akurasi random forest Model I (dengan pemilihan atribut)

Index Variabel	nTree	Training	Testing	Wacc Testing
1,2	50	0,9311	0,7238	0,287
1,2	150	0,9316	0,7232	0,2921
1,2	250	0,9316	0,7226	0,2844

Selain dari segi confusion matrix, dari segi nilai ukuran performa lain seperti precision dan recall. Seperti dapat dilihat pada Tabel 6-24, yang menjadi perhatian khusus adalah kelas 3 dan kelas 4. Precision kelas 3 berkisar antara 0,05-0,1 sementara kelas 4 tidak dapat dihitung yang menunjukkan ketepatan pengklasifikasi yang buruk. Recall kelas 3 berkisar 0,3-0,06 dan kelas 4 bernilai 0,00 yang menunjukkan kelengkapan pengklasifikasi yang buruk. Pada model ini kelas 4 belum dapat diprediksi oleh model.

Tabel 6-24 Nilai precision dan recall random forest Model I (dengan pemilihan atribut)

Kelas	50		150		250	
	Recall	Precision	Recall	Precision	Recall	Precision
0	0,8683	0,867612	0,872	0,865	0,869	0,864
1	0,4578	0,128931	0,445	0,130	0,455	0,121
2	0,2048	0,197674	0,193	0,188	0,217	0,217
3	0,0667	0,111111	0,03333 3	0,05263 2	0,06666 7	0,111111
4	0	#DIV/0!	0	#DIV/0!	0	#DIV/0!

6.2.1.3 Model II - dengan 8 Variabel Awal dan 4 Variabel Turunan

Pembuatan model II, menggunakan 8 variabel awal dari dataset, antara lain suhu udara, titik embun, tekanan permukaan laut, visibilitas, kecepatan angin, suhu udara minimum dan suhu udara maksimum serta 4 variabel turunan yaitu selisih suhu udara dan titik embun, estimasi kelembapan relatif, selisih suhu udara maksimum dan minimum serta bulan.

Seperti dapat dilihat pada Tabel 6-25, nilai ukuran performa model ini untuk akurasi data training stabil sebesar 0,99-1, nilai akurasi testing bernilai sebesar berkisar antara 0,78 dan nilai weghted accuracy untuk data testing berkisar 0,3-0,31. Nilai weighted accuracy memiliki nilai kecil karena adanya ketidakseimbangan jumlah data untuk tiap kelas.

Tabel 6-25 Nilai akurasi random forest Model I (dengan pemilihan atribut)

nTree	Training	Testing	Wacc Testing
50	0,999	0,7871	0,3081
150	1	0,7847	0,313
250	1	0,7824	0,3087

Selain dari segi confusion matrix, dari segi nilai ukuran performa lain seperti precision dan recall. Seperti dapat dilihat pada Tabel 6-26, yang menjadi perhatian khusus adalah kelas 3 dan kelas 4. Precision kelas 3 tidak dapat dihitung dan kelas 4 juga tidak dapat dihitung yang menunjukkan ketepatan pengklasifikasi yang sangat buruk. Recall kelas 3 bernilai 0 dan kelas 4 bernilai 0 yang menunjukkan kelengkapan pengklasifikasi yang cukup buruk. Pada model ini kelas 3 dan 4 belum dapat diprediksi oleh model.

Tabel 6-26 Nilai precision dan recall (dengan pemilihan atribut)

Kelas	50		150		250	
	Recall	Precision	Recall	Precision	Recall	Precision
0	0,924	0,879	0,934	0,884	0,932	0,881
1	0,523	0,157	0,545	0,169	0,529	0,177
2	0,157	0,406	0,145	0,462	0,120	0,455
3	0	#DIV/0!	0	#DIV/0!	0	#DIV/0!
4	0	#DIV/0!	0	#DIV/0!	0	#DIV/0!

6.2.1.4 Model dengan Variabel hasil Pemilihan Variable Importance dari Model II

Menggunakan 5 variabel untuk menghasilkan model selanjutnya, dimana hanya menggunakan variabel yang memiliki nilai kepentingan di atas rata-rata, yaitu suhu udara, suhu udara maksimum, selisih suhu udara dan titik embun, selisih suhu udara maksimum dan minimum serta estimasi kelembapan relatif. Nilai kepentingan yang dipilih adalah yang lebih besar dari nilai rata-rata.

Tabel 6-27 Kepentingan variabel random forest Model II (8 variabel awal, 4 variabel baru)

Index	50	150	250
1	1,484	1,299	1,261
2	1,057	1,040	0,893
3	0,851	0,813	0,780
4	0,355	0,352	0,483
5	0,896	0,581	0,766
6	0,689	0,453	0,527
7	1,654	1,550	1,527
8	0,872	0,817	0,722

9	1,337	1,657	1,560
10	1,624	1,569	1,489
11	2,423	1,760	2,049
12	1,120	0,893	0,880

Mean tree 50 = 1,197

Mean tree 150 = 1,065

Mean tree 250 = 1,078

Variabel dipilih yang memiliki nilai kepentingan yang dipilih adalah yang lebih besar dari nilai rata-rata, yaitu berkisar antara 1-1,19 seperti dapat dilihat pada Tabel 6-27.

Seperti dapat dilihat pada Tabel 6-28, nilai ukuran performa model ini untuk akurasi data training berkisar antara 0,99 , nilai akurasi testing bernilai berkisar pada 0,77 dan nilai weighted accuracy untuk data testing berkisar antara 0,31-0,32. Nilai weighted accuracy memiliki nilai kecil karena adanya ketidakseimbangan jumlah data untuk tiap kelas. Semakin banak proporsi data yang tidak dapat diprediksi dengan benar, maka semakin kecil nilai weighted accuracy.

Tabel 6-28 Nilai akurasi random forest Model II (dengan pemilihan atribut)

Index Variabel	nTree	Training	Testing	Wacc Testing
1,7,9,10,11	50	0,999	0,7711	0,3214
1,7,9,10,11	150	0,9995	0,7735	0,3125
1,7,9,10,11	250	0,9995	0,7753	0,3151

Selain dari segi confusion matrix, dari segi nilai ukuran performa lain seperti precision dan recall. Seperti dapat dilihat pada Tabel 6-29, yang menjadi perhatian khusus adalah kelas 3 dan kelas 4. Precision kelas 3 berkisar antara 0-0,03 sementara kelas 4 tidak dapat dihitung yang menunjukkan ketepatan pengklasifikasi yang buruk. Recall kelas 3 berkisar 0-0,03 dan kelas 4 bernilai 0,00 yang menunjukkan

kelengkapan pengklasifikasi yang buruk. Pada model ini kelas 4 belum dapat diprediksi oleh model.

**Tabel 6-29 Nilai precision dan recall random forest Model II
(dengan pemilihan atribut)**

Kelas	50		150		250	
	Recall	Precision	Recall	Precision	Recall	Precision
0	0,9172	0,878399	0,917	0,875	0,920	0,875
1	0,4968	0,174847	0,484	0,169	0,481	0,171
2	0,1205	0,25641	0,133	0,275	0,133	0,275
3	0	0	0	0	0,033333	0,333333
4	0	#DIV/0!	0	#DIV/0!	0	#DIV/0!

6.2.1.5 Perbandingan Model Random Forest

Seperti dapat dilihat pada Tabel 6-30, Model I (dengan 8 variabel) memiliki nilai ukuran performa akurasi data training yang paling tinggi yaitu 1, akurasi data testing paling tinggi yaitu 0,78 dan nilai weighted accuracy 0,3-0,31 namun memiliki nilai precision dan recall untuk yang buruk karena tidak mampu memprediksi satu pun kelas 3 dan kelas 4 yang benar.

Tabel 6-30 Perbandingan nilai akurasi model random forest

Ukuran Performa	Model I (8 Variabel)	Model I dengan pemilihan Attribut	Model II (12 Variabel)	Model II dengan pemilihan Attribut
Akurasi Training	1	0,93	0,99-1	0,99
Akurasi Testing	0,77-0,78	0,72	0,78	0,77

Weighted Accuracy Testing	0,3-0,31	0,28-0,29	0,30-0,31	0,31-0,32
---------------------------	----------	-----------	-----------	-----------

Tabel 6-31 adalah rata-rata nilai precision dan recall untuk pohon 50,150 dan 250.

Tabel 6-31 Perbandingan rata-rata nilai precision dan recall model random forest

K E L A S	Model I (8 Variabel)		Model I dengan pemilihan Attribut		Model II (12 Variabel)		Model II dengan pemilihan Attribut	
	Recal I	Preci sion	Recall	Preci sion	Rec all	Preci sion	Rec all	Preci sion
0	0,935	0,866	0,870	0,866	0,930	0,881	0,918	0,876
1	0,479	0,176	0,453	0,127	0,532	0,168	0,487	0,172
2	0,072	0,254	0,205	0,201	0,141	0,441	0,129	0,269
3	0,000	#DIV/ 0!	0,056	0,092	0,000	#DIV/ 0!	0,011	0,111
4	0,935	0,866	0,870	0,866	0,930	0,881	0,918	0,876

Model yang terbaik dipilih adalah model II dengan pemilihan atribut (yaitu suhu udara,suhu udara maksimum,selisih suhu udara dan titik embun, selisih suhu udara maksimum dan minimum seta estimasi kelembapan relatif), karena memiliki nilai relatif lebih besar dibandingkan dengan model yang lain, untuk nilai akurasi training yaitu 0.99, untuk nilai akurasi testing paling tinggi yaitu 0.77, nilai weighted accuracy paling tinggi yaitu 0.31-0,32 serta mampu memprediksi kelas 0,1,2 dan 3 jika dilihat dari precision dan recall.

Berdasarkan variabel yang terpilih berdasarkan nilai variable importance, variabel yang paling berpengaruh antara lain suhu udara, titik embun, suhu udara maksimum, selisih suhu udara dan titik embun estimasi kelembapan relatif serta selisih suhu udara maksimum dan minimum yang menunjukkan 3 dari 4 variabel turunan yaitu selisih suhu udara dan titik embun, estimasi kelembapan relatif dan selisih suhu udara maksimum dan minimum, memiliki nilai signifikan terhadap peningkatan nilai performa model.

6.2.2 Model dengan Pengendalian Jumlah Pohon dan Jumlah Variabel Acak

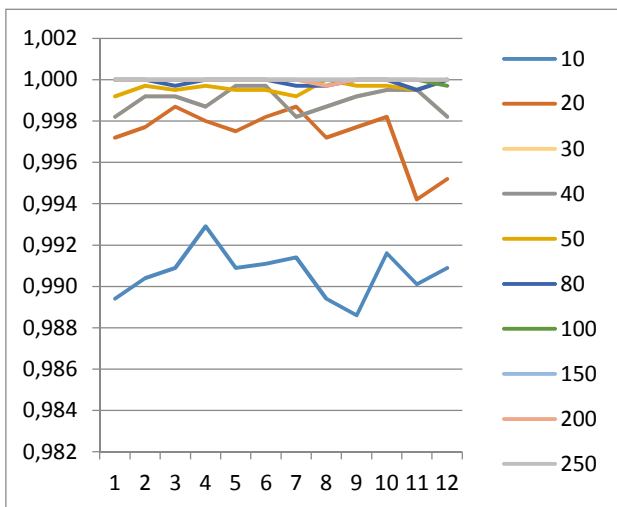
Pada tahap ini akan dilakukan analisa hasil pembuatan model dengan metode random forest dengan pengendalian jumlah pohon dan variabel acak yaitu melakukan tuning parameter jumlah pohon dan variabel acak.

6.2.2.1 Hasil Uji Performa Berdasarkan Akurasi Training, Akurasi Testing dan Weighted Accuracy

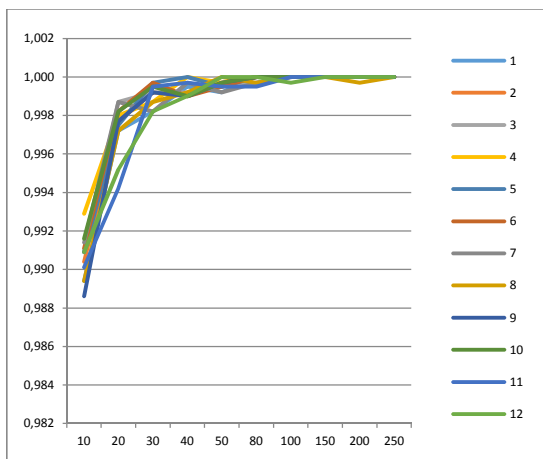
Pada tahap ini dianalisa hasil uji performa model dengan kombinasi jumlah pohon dan variabel acak dari segi nilai akurasi training, akurasi testing dan nilai weighted accuracy.

A. Akurasi data training

Pada bagian ini dijelaskan nilai performa akurasi data training. Pada Gambar 6-11 menunjukkan akurasi data training paling tinggi dan stabil dengan jumlah pohon >10 untuk semua kemungkinan jumlah variabel acak dengan nilai akurasi training sebesar 0.996-1.



Gambar 6-11 Grafik perbandingan jumlah variabel acak (x) terhadap jumlah pohon (y) Random Forest untuk akurasi training

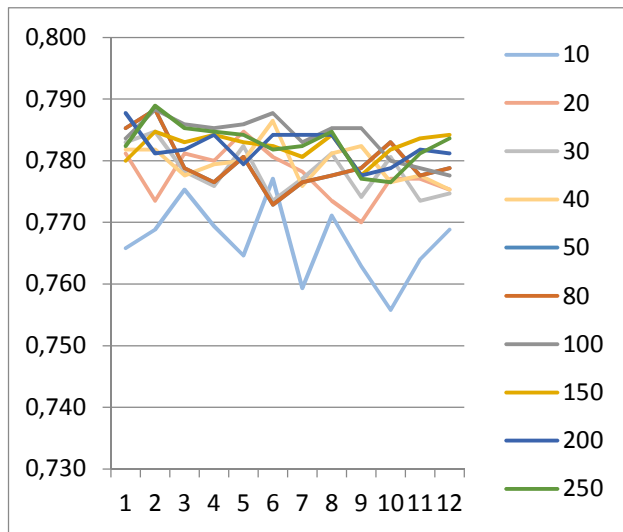


Gambar 6-12 Grafik perbandingan jumlah jumlah pohon (x) terhadap variabel acak (y) Random Forest untuk akurasi training

Pada Gambar 6-12 menunjukkan akurasi data training paling tinggi dan stabil dengan jumlah pohon >50 untuk semua kemungkinan jumlah variabel acak dengan nilai akurasi training sebesar 0.998-1. Nilai akurasi training lebih detail dapat dilihat pada Tabel D-1.

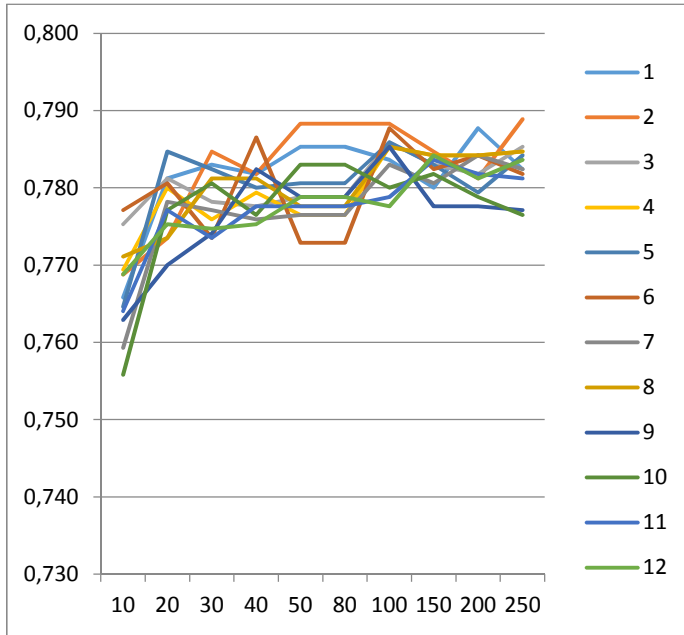
B. Akurasi data Testing

Pada bagian ini dijelaskan nilai performa akurasi data testing.



Gambar 6-13 Grafik perbandingan jumlah variabel acak (x) terhadap jumlah pohon (y) Random Forest untuk data testing

Pada Gambar 6-13 menunjukkan akurasi data testing paling tinggi dengan jumlah variabel acak 1-2 dengan nilai akurasi testing sebesar 0,79 kemudian menurun hingga sekitar 0,78 untuk variabel acak senilai 12, dalam hal ini kurang begitu ada penurunan signifikan. Stabil pada jumlah variabel acak 10-12.

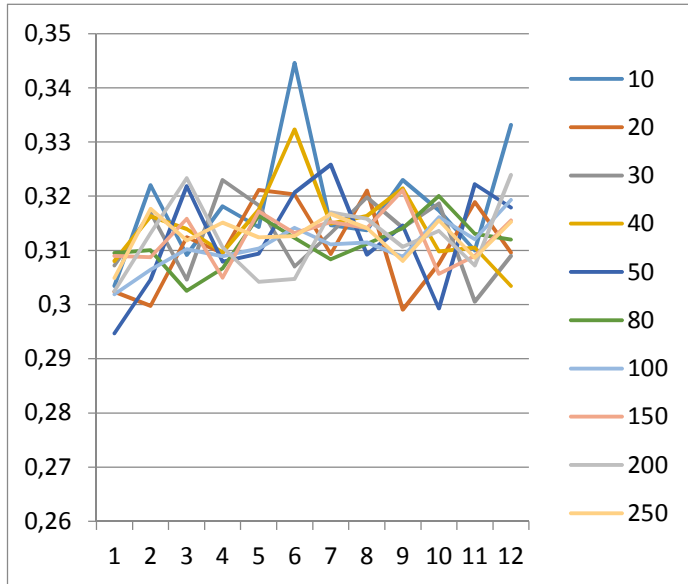


Gambar 6-14 Grafik perbandingan jumlah jumlah pohon (x) terhadap variabel acak (y) Random Forest untuk data testing

Pada Gambar 6-14 menunjukkan akurasi data training paling tinggi dan stabil dengan jumlah pohon >50 untuk semua kemungkinan jumlah variabel acak dengan nilai akurasi training sebesar 0.78-0.79. Nilai akurasi testing lebih detail dapat dilihat pada Tabel D-2.

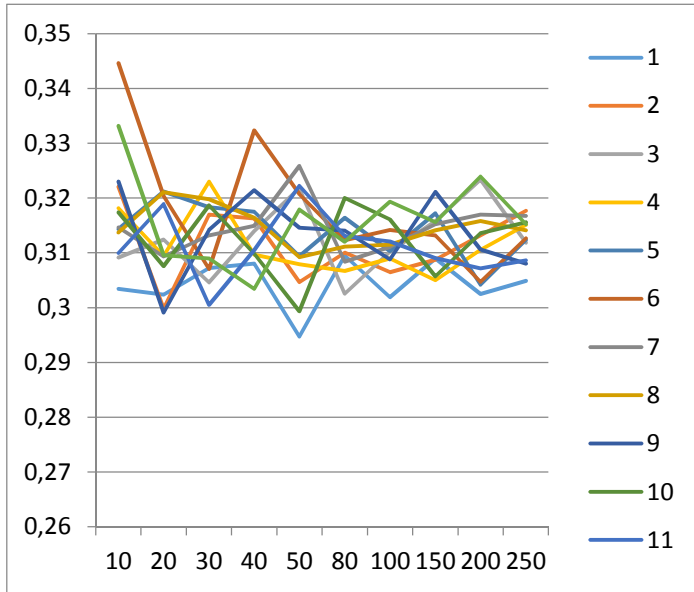
C. Weighted Accuracy data Testing

Pada bagian ini dijelaskan nilai perfroma weighted akurasi data testing.



Gambar 6-15) Grafik perbandingan jumlah variabel acak (x) terhadap jumlah pohon (y) Random Forest untuk nilai weighted accuracy data testing

Pada Gambar 6-15 menunjukkan weighted akurasi data testing paling tinggi dengan jumlah variabel 6 dengan nilai weighted akurasi testing mencapai 0,34. Nilai stabil pada jumlah variabel acak 7-10



Gambar 6-16 Grafik perbandingan jumlah jumlah pohon (x) terhadap variabel acak (y) Random Forest untuk nilai weighted accuracy data testing

Pada Gambar 6-16 menunjukkan akurasi paling stabil dengan jumlah pohon >20 untuk semua kemungkinan jumlah variabel acak dengan nilai akurasi training sebesar 0,3-0,32. Nilai weighted accuracy lebih detail dapat dilihat pada Tabel D-3. Nilai paling tinggi dengan jumlah pohon 10.

6.2.2.2 Hasil Uji Performa Berdasarkan Presisi dan Recall

Pada tahap ini dianalisa hasil uji performa model dengan kombinasi jumlah variabel acak dan jumlah pohon dari segi nilai nilai presisi dan recall.

A. Presisi Model untuk Kelas 0,1,2

Pada kelas 0,1,2 semua model uji dapat mengklasifikasi dengan baik, atau paling tidak terdapat kelas yang diprediksi

dengan benar. Nilai presisi model kelas 0,1,2 lebih detail dapat dilihat pada Tabel D-4, Tabel D-5, Tabel D-6.

B. PRESISI Model untuk Kelas 3

Seperti dapat dilihat pada Tabel 6-32, model yang mampu memprediksi kelas 3 dengan benar dan stabil hanya model dengan jumlah pohon 10 dan 20 untuk semua jumlah variabel acak. Selain model tersebut, model mulai tidak stabil memprediksi kelas 3, yang ditunjukkan dengan NaN yaitu nilai yang dihasilkan karena nilai prediksi benar sama dengan 0 yang dibagi dengan total kelas prediksi yang juga bernilai 0.

Tabel 6-32 Nilai presisi model random fores (Pengendalian jumlah pohon dan variabel acak) kelas 3

	JUMLAH POHON									
Var. acak	10	20	30	40	50	80	100	150	200	250
1	0,00	0,00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	0,17	0,00	NaN	NaN	0,00	NaN	NaN	NaN	0,00	NaN
3	0,40	0,50	NaN	0,00	0,50	0,00	0,00	NaN	NaN	NaN
4	0,17	0,00	0,00	0,00	NaN	NaN	NaN	0,00	0,00	0,00
5	0,17	NaN	1,00	NaN	0,00	NaN	1,00	NaN	NaN	NaN
6	0,33	0,50	NaN	NaN	NaN	0,00	NaN	NaN	NaN	NaN
7	0,50	0,33	0,00	0,00	0,00	0,00	NaN	0,00	0,33	NaN
8	0,20	0,50	0,50	0,00	0,00	NaN	0,50	NaN	1,00	1,00
9	0,38	0,50	0,50	1,00	0,00	0,00	1,00	NaN	NaN	NaN
10	0,09	0,00	1,00	0,00	0,33	1,00	0,00	NaN	NaN	NaN
11	0,43	0,33	0,00	NaN	0,50	0,00	NaN	1,00	0,50	NaN
12	0,17	0,40	0,33	0,00	0,00	NaN	1,00	0,00	NaN	NaN

C. PRESISI Model untuk Kelas 4

Semua model tidak mampu memprediksi kelas 4 dengan benar dan stabil, yang ditunjukkan dengan NaN yaitu nilai yang dihasilkan karena nilai prediksi benar sama dengan 0 yang dibagi dengan total kelas prediksi yang juga bernilai 0. Nilai presisi model kelas 4 lebih detail dapat dilihat pada Tabel D-7.

D. Recall Model untuk Kelas 0,1,2

Pada kelas 0,1,2 semua model uji dapat mengklasifikasi dengan baik, atau paling tidak terdapat kelas yang diprediksi dengan benar. Nilai recall kelas 0,1,2 lebih detail dapat dilihat pada Tabel D-8, Tabel D-9 dan Tabel D-10.

E. RECALL Model untuk Kelas 3

Seperti dapat dilihat pada Tabel 6-33, model yang mampu memprediksi kelas 3 dengan benar dan stabil hanya model dengan jumlah pohon 10 dan 20 untuk semua jumlah variabel acak. Selain model tersebut, model mulai tidak stabil memprediksi kelas 3, yang ditunjukkan dengan mulai banyaknya nilai 0 yaitu nilai yang dihasilkan karena nilai prediksi benar sama dengan 0 yang dibagi dengan total kelas prediksi yang bernilai >1 .

Tabel 6-33 Nilai recall model random fores (Pengendalian jumlah pohon dan variabel acak) kelas 3

	JUMLAH POHON									
Var. acak	10	20	30	40	50	80	100	150	200	250
1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
2	0,07	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00
3	0,00	0,07	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00
4	0,03	0,03	0,00	0,03	0,03	0,03	0,00	0,00	0,00	0,00
5	0,03	0,07	0,00	0,03	0,00	0,00	0,00	0,03	0,00	0,00

	JUMLAH POHON									
Var. acak	10	20	30	40	50	80	100	150	200	250
6	0,10	0,07	0,00	0,03	0,03	0,00	0,03	0,00	0,00	0,03
7	0,07	0,07	0,00	0,03	0,03	0,03	0,00	0,00	0,03	0,00
8	0,00	0,00	0,07	0,00	0,03	0,00	0,00	0,03	0,03	0,00
9	0,07	0,07	0,03	0,00	0,00	0,03	0,00	0,00	0,03	0,00
10	0,03	0,03	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,03
11	0,00	0,07	0,00	0,07	0,00	0,00	0,00	0,00	0,00	0,03
12	0,10	0,00	0,00	0,00	0,00	0,03	0,03	0,00	0,03	0,00

F. RECALL Model untuk Kelas 4

Semua model tidak mampu memprediksi kelas 4 dengan benar dan stabil, yang ditunjukkan dengan 0 yaitu nilai yang dihasilkan karena nilai prediksi benar sama dengan 0 yang dibagi dengan total kelas prediksi yang bernilai >1. Nilai recall kelas 4 lebih detail dapat dilihat pada Tabel D-11.

6.2.3 Pengujian Model Random Forest dengan Proporsi Presentase Data Training dan Testing

Pada Tabel 6-34 menunjukkan untuk semua kombinasi proporsi data training dan testing untuk model paling baik dari tahap pengendalian besar random forest, dengan jumlah atribut acak 6 dan jumlah pohon 10, menghasilkan akurasi yang cenderung stabil, yaitu untuk data training berkisar antara 0,98-0,99 dan untuk data testing berkisar antara 0,76-0,78 dengan 0,789 dimiliki oleh proporsi data training-testing sebesar 60%-40%. Pada nilai Weighted Accuracy, cenderung stabil berkisar antara 0,29-0,33.

Tabel 6-34 Nilai akurasi pengujian proporsi data untuk random forest

Training	Testing	Tr	Test	WaccT
60%	40%	0,989	0,789	0,322
70%	30%	0,991	0,788	0,330
75%	25%	0,990	0,780	0,330
80%	20%	0,991	0,766	0,307
85%	15%	0,986	0,764	0,296
90%	10%	0,990	0,769	0,326

6.2.4 Kesimpulan Hasil Model Random Forest

Berikut ini adalah kesimpulan model Random Forest:

1. Variabel yang dipilih untuk dilibatkan dalam pembuatan model adalah model II dengan pemilihan atribut (5 variabel: suhu udara, suhu udara maksimum, selisih suhu udara dan titik embun, selisih suhu udara maksimum dan minimum serta estimasi kelembapan relatif) karena memiliki nilai uji performa yang baik dari segi akurasi dan dari segi presisi dan recall. Model ini menggunakan parameter default jumlah tree berkisar 50-250 dan jumlah variabel acak sejumlah 2 dan proporsi training/testing 70%/30%.
2. Model II dengan pemilihan atribut (5 variabel) memiliki akurasi training yaitu 0.99, untuk nilai akurasi testing paling tinggi yaitu 0.77, nilai weighted accuracy paling tinggi yaitu 0.31-0.32 serta mampu memprediksi kelas 0,1,2 dan 3 jika dilihat dari precision dan recall.
3. Pengendalian jumlah pohon dan jumlah variabel acak: berikut ini adalah hasil akurasi sesuai hasil masing-masing uji performa, dibagi menjadi 2 bagian yaitu hasil yang stabil dan yang tinggi, seperti dapat dilihat pada Tabel 6-35:

**Tabel 6-35 Nilai parameter berdasar kan uji performa
pengendalian jumlah pohon dan jumlah variabel model random
forest**

	Train	Test	WaccTest	Presisi (memprediksi 0,1,2,3)	Recall (memprediksi 0,1,2,3)
<i>Stabil</i>					
<i>Variabel Acak</i>	All	10-12	7-10	All	All
<i>Pohon</i>	50-250	50-250	20-250	10-20	10-20
<i>Tinggi</i>					
<i>Variabel Acak</i>	All	2	6		
<i>Pohon</i>	50-250	50-250	10		

Sehingga, model yang paling baik yang dapat memprediksi 0,1,2,3 adalah dengan jumlah variabel acak 6 dan jumlah pohon (10). Model tersebut pada sisi presisi (ketepatan pengklasifikasi) untuk kelas 0 memiliki performa 0.875, kelas 1 sebesar 0.471, kelas 2 sebesar 0.392, kelas 3 sebesar 0.33 dan kelas 4 tidak bisa diklasifikasikan dengan benar.. Dari segi recall (kelengkapan pengklasifikasi) untuk kelas 0 memiliki performas sebesar 0.906, kelas 1 sebesar 0.516, kelas 2 sebesar 0.198 , kelas 3 sebesar 0.1 dan kelas 4 tidak berhasil diklasifikasi. Hal ini menunjukkan kemampuan pengklasifikasi model paling baik pada kelas 0, cukup baik pada kelas 1 (hanya sekitar 50% benar) , pada kelas 2 dan 3 hanya berkisar 10-20% yang benar.

- Proporsi training dan testing untuk model Random Forest dengan jumlah variabel acak 6 dan jumlah pohon 10 untuk proporsi 60:40, 70:30, 75:25, 80:20, 85:15 dan 90:10 cenderung stabil, untuk data training berkisar antara 0,98-0,99 dan untuk data testing

berkisar antara 0,76-0,78 serta nilai Weighted Accuracy, cenderung stabil berkisar antara 0,29-0,33.

5. Melakukan pengujian kestabilan proporsi data training/testing untuk model Random Forest dengan jumlah variabel acak 6 dan jumlah pohon 10 menunjukkan proporsi 70% data training dan 30% data testing dengan nilai akurasi training 0,991, akurasi testing 0,788 dan weighted accuracy testing 0,330.

6.3 Analisa Perbandingan Hasil CART dan Random Forest

1. Pada Tabel 6-36 adalah analisa perbandingan dari segi kepentingan variabel, untuk kedua metode:

Tabel 6-36 Perbandingan nilai uji performa CART dan Random forest

CART				
Ukuran Performa	Model I (8 Variabel)	Model I dengan pemilihan Attribut [1,2]	Model II (12 Variabel)	Model II dengan pemilihan Attribut [9,11]
Akurasi Training	0.9204	0.8596	0.9321	0.8781
Akurasi Testing	0.7297	0.7363	0.7286	0.7280
Weighted Accuracy Testing	0.3015	0.2856	0.3216	0.3146
RANDOM FOREST				
Ukuran Performa	Model I (8 Variabel)	Model I dengan pemilihan Attribut [1,2]	Model II (12 Variabel)	Model II dengan pemilihan Attribut [1,7,9,10,11]
Akurasi Training	1	0,93	0,99-1	0,99
Akurasi	0,77-0,78	0,72	0,78	0,77

Testing				
Weighted Accuracy Testing	0,3-0,31	0,28-0,29	0,30-0,31	0,31-0,32

Menggunakan parameter default, terdapat perbedaan 6% tingkat akurasi antara CART dan Random Forest, yaitu untuk CART berkisar antara 0,72-0,73 sedangkan Random Forest adalah 0,78. Parameter default yang dimaksud adalah yang menurut Breiman [17], untuk CART dengan minimal leaf node 10 dan minimal parent 1, sedangkan untuk Random Forest jumlah variabel acak akar jumlah variabel.

Pada CART, pengurangan jumlah variabel masih mampu mempertahankan nilai akurasi jika dibandingkan model dengan variabel penuh, yang menandakan bahwa variabel tersebut memiliki kontribusi besar pada pembangunan model. Variabel yang memiliki kontribusi besar tersebut antara lain index 1,2,9 dan 11 yaitu suhu udara, titik embun, selisih suhu udara dan titik embun serta selisih suhu udara minimum dan maksimum. Sehingga, dapat ditarik kesimpulan untuk model CART, variabel yang berpengaruh besar dalam pembuatan model adalah suhu udara, titik embun, suhu udara maksimum, suhu udara minimum dan variabel turunannya.

Pada Random Forest variabel yang berkontribusi besar antara lain index 1,2,7,9,10,11 yaitu suhu udara, titik embun, suhu udara maksimum, selisih suhu udara dan titik embun, kelembapan relatif dan selisih suhu udara maksimum dan minimum. Dari kedua metode, dapat ditarik kesimpulan bahwa suhu udara, titik embun, suhu udara maksimum dan suhu udara minimum beserta variabel turunannya (selisih suhu udara maksimum dan minimum, selisih suhu udara dan titik embun dan kelembapan relatif) yang berpengaruh besar terhadap pembangunan model, dengan tingkat akurasi mencapai 0,78.

2. Pada Tabel 6-37 analisis kedua metode dari segi presisi dan recall,

Tabel 6-37 Perbandingan nilai precision dan recall CART dan random forest

K E L A S	CART							
	Model I (8 Variabel)		Model I dengan pemilihan Attribut		Model II (12 Variabel)		Model II dengan pemilihan Attribut	
	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
0	0,868	0,862	0,883	0,856	0,861	0,870	0,863	0,860
1	0,432	0,162	0,416	0,175	0,445	0,176	0,448	0,160
2	0,139	0,192	0,129	0,250	0,129	0,186	0,089	0,176
3	0,069	0,071	0,000	0,000	0,172	0,192	0,172	0,167
	0,000	#DIV/0!	0,000	#DIV/0!	0,000	#DIV/0!	0,000	#DIV/0!
K E L A S	RANDOM FOREST							
	Model I (8 Variabel)		Model I dengan pemilihan Attribut		Model II (12 Variabel)		Model II dengan pemilihan Attribut	
	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
0	0,935	0,866	0,870	0,866	0,930	0,881	0,918	0,876
1	0,479	0,176	0,45	0,12	0,53	0,16	0,48	0,1

			3	7	2	8	7	72
2	0,072	0,254	0,20 5	0,20 1	0,14 1	0,44 1	0,12 9	0,2 69
3	0,000	#DIV/0!	0,05 6	0,09 2	0,00 0	#DIV /0!	0,01 1	0,1 11
4	0,000	#DIV/0!	0,00 0	#DIV /0!	0,00 0	#DIV /0!	0,00 0	#DI V/0 !

Menggunakan 2 variabel (suhu udara dan titik embun), CART hanya mampu mengklasifikasi kelas 0,1,2 sedangkan untuk Random Forest mampu mengklasifikasi kelas 0,1,2,3 dengan nilai akurasi testing yang relatif sama yaitu 0,72 untuk random forest dan 0,736 untuk CART.

Pada Random forest bila melibatkan semua variabel (tanpa memperhatikan kepentingan variabel), hanya mampu mengklasifikasi kelas 0,1,2 namun menghasilkan akurasi yang tinggi karena model juga mempertimbangkan variabel yang kurang mampu mengklasifikasi untuk semua kelas.

Pada CART, telah mampu mengklasifikasi kelas 0,1,2,3 namun memiliki nilai akurasi 6% lebih kecil dibandingkan random forest untuk parameter default.

3. Pada Tabel 6-38 adalah uji coba tuning parameter untuk kedua metode untuk mencari kombinasi parameter untuk menghasilkan model terbaik:

Tabel 6-38 Nilai parameter berdasarkan uji performa percobaan parameter model CART dan random forest

CART					
	Tra in	Tes t	Wacc Tes t	Presisi (memprediksi 0,1,2,3)	Recall (memprediks i 0,1,2,3)
<i>Stabil</i>					
<i>Leaf</i>	20- 100	30- 60	50-90		
<i>Pare</i>	50-	50-	All		

nt	90	90			
Tinggi					
Leaf	1-5	30-60	1-5	3-5	1-5
Parent	1-10	50-90	10	1-50	1-50
RANDOM FOREST					
	Tra in	Tes t	Wacc Tes t	Presisi (mempred iksi 0,1,2,3)	Recall (mempred iksi 0,1,2,3)
Stabil					
Variabel Acak	All	10-12	7-10	All	All
Pohon	50-250	50-250	20-250	10-20	10-20
Tinggi					
Variabel Acak	All	2	6		
Pohon	50-250	50-250	10		

Untuk CART, kombinasi parameter yang menghasilkan model terbaik dengan nilai akurasi yang tinggi yang dapat mengklasifikasi 0,1,2,3 adalah dengan leaf (1-5) dan parent (30). Hal ini berarti untuk CART maksimal, model mampu bekerja baik bila dalam 1 leafnode minimal memiliki 1-5 hasil observasi untuk menentukan label, yang berarti semakin sedikit nilai minimal leaf node maka akan semakin besar kedalaman pohon keputusan.

Untuk random forest, kombinasi parameter yang menghasilkan model terbaik dengan nilai akurasi yang tinggi adalah yang dapat mengklasifikasi 0,1,2,3 adalah dengan jumlah variabel acak 6 dan jumlah pohon (10). Hal ini berarti dalam pembangunan model random forest, tidak memerlukan begitu banyak pohon namun memerlukan jumlah variabel acak lebih banyak dari parameter default.

4. Pada Tabel 6-39 adalah hasil analisis percobaan proporsi data training dan testing untuk model terbaik (dari poin 3) dari masing-masing metode:

Tabel 6-39 Hasil nilai akurasi percobaan proporsi data untuk CART dan random forest

CART				
Training	Testing	Tr	Test	WaccT
60%	40%	0,858	0,775	0,319
70%	30%	0,857	0,767	0,324
75%	25%	0,855	0,768	0,323
80%	20%	0,861	0,748	0,323
85%	15%	0,862	0,763	0,320
90%	10%	0,858	0,761	0,300
RANDOM FOREST				
Training	Testing	Tr	Test	WaccT
60%	40%	0,989	0,789	0,322
70%	30%	0,991	0,788	0,330
75%	25%	0,990	0,780	0,330
80%	20%	0,991	0,766	0,307
85%	15%	0,986	0,764	0,296
90%	10%	0,990	0,769	0,326

Baik CART, maupun Random forest memiliki kestabilan dari segi nilai akurasi untuk beberapa kombinasi proporsi data training dan data testing, yaitu 60:40, 70:30, 75:25, 80:20, 85:15 dan 90:10.

Halaman ini sengaja dikosongkan

BAB VII

KESIMPULAN DAN SARAN

Pada bab ini dibahas mengenai kesimpulan dari semua proses yang telah dilakukan dan saran yang dapat diberikan untuk pengembangan yang lebih baik.

7.1 Kesimpulan

Berikut ini adalah kesimpulan dari hasil penelitian tugas akhir:

1. Model dengan menggunakan parameter default sesuai Breiman, baik untuk metode CART dan Random Forest untuk dataset cuaca yang digunakan, berdasarkan hasil performa menunjukkan model Random Forest sedikit lebih baik di atas CART sebesar 6%.
2. Untuk data dengan tingkat ketidakseimbangan yang tinggi antar kelas, seperti dataset yang digunakan dalam tugas akhir ini, baik model CART maupun dan Random Forest dengan tuning parameter yang optimal memiliki kemampuan yang cukup baik yaitu mampu menghasilkan nilai akurasi sebesar 77%-78% dan mampu mengklasifikasi 4 kelas (0,1,2,3) dari 5 kelas (0,1,2,3,4). Kelas 4 (kelas hujan sangat lebat) tidak mampu diklasifikasi oleh kedua metode dengan benar meskipun telah dilakukan tuning parameter (pemilihan parameter) untuk mencapai model terbaik. Performa model dari segi presisi dan recall menunjukkan untuk model dengan tuning parameter memiliki performa akurasi paling baik pada kelas 0 (80%-90%), kelas 1 (40%-50%), kelas 2 dan 3 (10%-20%) sedangkan kelas 4 sebesar 0%.
3. Model CART maupun dan Random Forest dengan tuning parameter yang optimal memiliki kestabilan nilai performa untuk proporsi data training/testing sebesar 60:40, 70:30, 75:25, 80:20, 85:15 dan 90:10.

4. Variabel yang berpengaruh besar dalam konstruksi model CART dan random forest dengan nilai uji performa yang baik berdasarkan nilai *variable importance* antara lain adalah suhu udara, titik embun, suhu udara maksimum dan suhu udara minimum beserta variabel turunannya (selisih suhu udara maksimum dan minimum, selisih suhu udara dan titik embun dan kelembapan relatif), dengan tingkat akurasi mencapai relatif baik. Sehingga, dapat disimpulkan bahwa 3 dari 4 variabel turunan yaitu selisih suhu udara dan titik embun, estimasi kelembapan relatif dan selisih suhu udara maksimum dan minimum, bernilai signifikan terhadap peningkatan nilai performa model.

7.2 Saran

Berikut ini adalah beberapa saran pengembangan untuk penelitian tugas akhir:

1. Menggunakan Analisis *outlier* untuk melihat karakteristik dan mendeteksi kelas 4 (hujan sangat lebat) karena kelas 4 merupakan data outlier, yaitu 7 data curah hujan yang sangat lebat dari 5635 data curah hujan.
2. Menggunakan data cuaca dari 2 stasiun cuaca lain di Surabaya, yaitu stasiun Perak dan stasiun Gedangan.
3. Menambahkan variabel/attribut baru seperti kelembapan rata-rata, kelembapan maksimum dan kelembapan minimum.

DAFTAR PUSTAKA

- [1] &. S. K. Baboo S., "Applicability of Data Mining Techniques for Climate Prediction – A Survey Approach," *International Journal of Computer Science and Information Security*, vol. 8, April 2010.
- [2] W. W. H. Y. Zhang Z., "Mining dynamic interdimension association rules for local-scale weather prediction," *In the Proceedings of the 28th Annual International Computer software and Applications Conference*, 2004.
- [3] K. P. RI, "Peraturan Menteri Perhubungan Republik Indonesia no 9 tahun 2015 tentang Peraturan Keselamatan Penerbangan Sipil Bagian 174 Tentang Pelayanan Informasi Meteorologi Penerbangan (Aeronautical Meteorological Information Services," *Peraturan Menteri Perhubungan Republik Indonesia no 9 tahun 2015 tentang Peraturan Keselamatan Penerbangan Sipil Bagian 174 (ivil Aviation Safety Regulations Part 174) Tentang*, 2015.
- [4] T. Z. Tan, "Rainfall intensity prediction by a spatial-temporal ensemble," *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence*, 2008.
- [5] T. R. Prajwala, "A Comparative Study on Decision Tree and Random Forest Using R Tool," vol. 4, no. 1, 2015.
- [6] C. D. Sutton, "Classification and Regression Trees, Bagging, and Boosting," *Handbook of Statistics*, 2005.
- [7] D. A. A. C. J. K. T. R. S. M. S. a. S. D. John K. Williams, "A Machine Learning Approach to Finding Weather Regimes and Skillful Predictor Combinations for Short-term Storm Forecasting," *National Center for Atmospheric Research*, 2008.
- [8] E. G. Petre, "A Decision Tree for Weather Prediction," *Buletinul UniversităŃii Petrol – Gaze din Ploieşti*, 2009.

- [9] K. d. G. Badan Meteorologi, September 2016. [Online]. Available: <http://meteo.bmkg.go.id/prakiraan/mingguan>.
- [10] Institute Indian Agricultural Statistics Research, Data Preprocessing Techniques for Data Mining, Institute Indian Agricultural Statistics Research, 2007.
- [11] D. B. Rubin, "Inference and missing data," *Oxford Hournal Biometrika*, 1975.
- [12] S. U. R. J. a. J. J. M. Timothy Hayes, "Using Classification and Regression Trees (CART) and Random Forests to Analyze Attrition: Results From Two Simulations," *Psychol Aging*, 2015.
- [13] R. K. a. M. S. J. Dougherty, "Supervised and unsupervised discretization of continuous features," *In Machine Learning: Proceedings of the Twelfth International Conference*, 1995.
- [14] A. Ahmad, "Data Transformation For Decision Tree Ensembles," *Thesis for Degree of Doctor of Phylosophy University Of Manchester*, 2009.
- [15] M. J. d. L. G. S. Berry, Data Mining Techniques For Marketing, Sales, Customer Relationship Management Second Editon, United States of America: Wiley Publishing, Inc , 2004.
- [16] S. Soni, "Implementation of Multivariate Data Set By Cart Algorithm," *International Journal of Information Technology and Knowledge Management*, 2010.
- [17] F. J. O. R. S. C. Breiman L, Classification and Regression Trees, New York: Chapman & Hall, 1993.
- [18] F. J. O. R. S. C. Brieman L, Classification and Regression Trees, New York: Chapman & Hall, 1984.
- [19] S. J. J. K. u. I. R. K. Anne-Laure Boulesteix, "Over view of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics," *Technical Report Number 129, Department of Statistics, University of Munich*, 2012.

- [20] M. Zakariah, "Classification of genome data using Random Forest Algorithm: Review," *Interational Journal Computer Technology and Application (IJCTA)*, 2014.
- [21] E. C. P. a. F. B. S. B. B. A. Goldstein, "Random forests for genetic association studies," *Statistical Applications in Genetics and Molecular Biology*, 2011.
- [22] Mathworks, 2016. [Online]. Available: <http://www.mathworks.com/help/stats/fitctree.html>.
- [23] Mathworks, 2016. [Online]. Available: <http://www.mathworks.com/help/stats/treebagger.html>.
- [24] L. C. Thomas, "A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers.," *International Journal of Forecasting*, vol. 16(2), 2000.
- [25] O. A. a. R. E. E. Alduchov, " Improved Magnus' form approximation of saturation vapor pressure," *J. Appl. Meteor*, pp. 601-609, 1996.
- [26] Microsoft, Januari 2015. [Online]. Available: <https://msdn.microsoft.com/en-us/library/bb895173.aspx>.
- [27] PT PLN (Persero) Distribusi Jawa Timur, "Statistik 2012 - PT PLN (Persero) Distribusi Jawa Timur," Surabaya, 2012.
- [28] P. P. (. D. J. Timur, 2016.
- [29] D. Nachrowi and H. Usman, Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan, Jakarta: Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia, 2006.
- [30] J. G. Caldwell, 2006. [Online]. Available: <http://www.foundationwebsite.org/>.
- [31] C. Bennett, M. Moghimi, M. J. Hossain, J. Lu and R. A Stewart, "Applicability of Load Forecasting Techniques for Customer Energy Storage Control Systems," in *IEEE PES Asia-Pacific Power And Energy Engineering Conference (APPEEC)*, Brisbane, 2015.

- [32] J. Hassan, "ARIMA and Regression Models for Predicting of Daily and Monthly Clearness Index," *ScienceDirect*, pp. 421-427, March 2014.
- [33] P. Narayanan, A. Basistha, S. Sarkar and S. Kamna, "Trend Analysis and ARIMA Modelling of Pre-Monsoon Rainfall Data for Western India," *Elsevier Masson SAS*, pp. 22-27, December 2013.
- [34] K. Soni, S. Kapoor, K. Singh Pamar and D. G Kaskaoutis, "Statistical Analysis of Aerosol Over The Gangetic-Himalayan region Using ARIMA Model Based On Long-Term MODIS Observations," *Elsevier Atmospheric Research* 149, pp. 174-192, June 2014.
- [35] A. Herwindyani, "Pembuatan Aplikasi Pendukung Keputusan Untuk Peramalan Persediaan Bahan Baku Produksi Plastik Blowing dan Inject Menggunakan Metode ARIMA (Autoregressive Integrated mmoving Average) di CV Asia," 2014.
- [36] D. C. Montgomery, C. L. Jennings and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, New Jersey: John Wiley & Sons, Inc, 2008.
- [37] P.-C. Chang, Y.-W. Wang and C.-H. Liu, "The Development of a Weighted Ecolcing Fuzzy Neural Network for PCB Sales Forecasting," *Expert System with Application* 32, pp. 86-96, 2007.
- [38] S. Makridakis, S. C. Wheelwright and R. J. Hyndman, *Forecasting: Methods and Applications*, 3rd ed., John Wiley & Sons, 1998.
- [39] PT PLN (Persero), "Profil Perusahaan," 2011. [Online]. Available: <http://www.pln.co.id/blog/profil-perusahaan>. [Accessed 24 December 2015].
- [40] PT PLN (Persero) Distribusi Jawa Timur, "Golongan Tarif Dasar Listrik," 2009. [Online]. Available: <http://www.pln-jatim.co.id/red/?m=tdl2>. [Accessed 24 December 2015].
- [41] R. Weron, *Modeling and Forecasting Electricity Loads*

and Prices, West Sussex: John Wiley & Sons Ltd., 2006.

- [42] The MathWorks, Inc., "Box-Jenkins Methodology," [Online]. Available: <http://www.mathworks.com/help/econ/box-jenkins-methodology.html>. [Accessed 10 January 2016].
- [43] J. E. Hanke and D. W. Wichern, Business Forecasting, 9th ed., Prentice Hall, 2009.
- [44] O. Bowerman and K. , Forecasting, Time Series, and Regression, 4th ed., Curt Hinrichs, 2005.
- [45] L. Pham, "Time Series Analysis with ARIMA - ARCH/GARCH Model in R," 2013.
- [46] Minitab Inc., "Minitab StatGuide," 2007.
- [47] December 2014. [Online]. Available: <http://stats.stackexchange.com/questions/125926/how-to-read-unit-root-test-results-obtained-from-evIEWS-i-mean-what-values-do-w>.
- [48] R. J Hyndman, January 2014. [Online]. Available: <http://robjhyndman.com/hyndsight/ljung-box-test/>.
- [49] Microsoft, 2016. [Online]. Available: <https://support.office.com/en-us/article/EXP-function-c578f034-2c45-4c37-bc8c-329660a63abe>.
- [50] Eurostat, [Online]. Available: <https://ec.europa.eu/eurostat/sa-elearning/additive-decomposition-0>.
- [51] MediaWiki, January 2016. [Online]. Available: https://id.wikipedia.org/wiki/Fungsi_eksponensial.
- [52] D. M. Lane. [Online]. Available: <http://onlinestatbook.com/2/transformations/log.html>.

Halaman ini sengaja dikosongkan

BIODATA PENULIS



Penulis lahir di Surabaya, 04 Agustus 1996, dengan nama lengkap Maulana Dhawangkara. Penulis merupakan anak terakhir dari tiga bersaudara.

Riwayat pendidikan penulis yaitu TK Ananda, SD Negeri Banjarsugihan V Surabaya, SMP Negeri 2 Surabaya, SMA Negeri 2 Surabaya, dan akhirnya menjadi salah satu mahasiswa Sistem Informasi angkatan 2013 melalui jalur SNMPTN Undangan dengan

NRP 5213-100-074.

Selama kuliah, penulis berfokus terhadap bidang akademik dan berhasil beberapa kali meraih predikat IPK angkatan paling tinggi. Penulis pernah menjadi staff ahli tim soal olimpiade IT dan Bisnis yang diselenggarakan Sistem Informasi ITS, IS EXPO. Penulis mengambil bidang minat Rekayasa Data dan Intelegensi Bisnis. Penulis dapat dihubungi melalui email iwangkara95@gmail.com

Halaman ini sengaja dikosongkan

LAMPIRAN A

DATA MENTAH (8 VARIABEL AWAL)

Tabel A-1 Data mentah (8 variabel dengan satuan Indonesia)

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20000101	28,50	23,83	1006,4	9,01	6,67	18,33	33,78	23,89
20000105	28,50	24,22	1008,1	9,66	3,89	11,11	32,00	24,39
20000106	28,39	23,94	1007,9	7,40	5,00	22,22	33,00	24,78
20000107	26,67	23,44	1008,0	7,72	5,19	33,52	33,28	23,00
20000108	24,89	23,17	1009,2	6,60	6,30	14,82	29,00	22,39
20000109	27,11	23,94	1009,1	7,72	5,56	12,96	32,72	23,11
20000110	26,39	23,89	1007,8	8,53	4,82	9,45	32,39	23,00
20000111	25,83	23,56	1008,0	7,56	5,37	14,82	30,00	23,28
20000112	25,22	22,83	1007,3	7,89	3,15	12,96	30,00	22,72
20000113	28,50	22,67	1006,6	9,66	9,82	14,82	32,78	22,39
20000116	27,33	23,56	1008,4	9,17	5,00	18,33	32,11	22,61
20000117	26,44	23,44	1008,4	7,08	5,93	18,33	32,61	23,00
20000120	27,00	23,22	1007,7	8,85	9,63	18,33	31,00	24,00
20000121	27,83	22,94	1008,5	9,82	15,93	25,93	31,28	22,78
20000122	27,78	22,28	1007,3	9,01	20,19	38,89	32,39	24,22
20000123	26,67	23,06	1008,0	8,53	12,96	27,78	31,39	24,50
20000126	25,89	22,94	1007,9	8,53	10,37	18,33	29,28	24,00
20000127	27,11	22,94	1008,4	7,89	14,82	25,93	30,78	23,72
20000130	25,94	23,06	1010,0	8,05	5,74	14,82	31,28	23,00
20000202	27,78	22,94	1008,6	9,01	13,33	18,33	32,00	23,61
20000203	27,00	23,06	1009,0	8,85	13,89	27,78	30,61	24,22
20000204	27,39	22,83	1008,3	9,17	12,78	22,22	30,22	25,00
20000205	28,11	23,22	1007,8	8,69	5,37	14,82	32,00	25,11
20000207	25,33	23,00	1006,9	6,28	7,78	18,33	29,22	23,50
20000208	25,83	23,61	1006,6	7,24	11,48	27,78	29,78	23,00
20000209	25,50	23,11	1007,5	7,08	10,00	22,22	30,28	23,22
20000210	26,78	22,83	1011,6	9,98	12,04	25,93	29,78	23,72
20000211	28,56	23,06	1008,5	9,66	13,33	18,33	32,00	24,72
20000212	28,33	23,56	1007,6	8,21	13,15	22,22	32,61	23,50
20000213	27,11	23,94	1007,0	6,28	3,70	11,11	31,61	22,61
20000215	27,06	24,17	1008,0	6,60	2,22	9,45	30,78	22,72
20000216	27,72	23,94	1007,5	8,53	6,85	18,33	33,00	22,78
20000217	28,56	24,00	1007,4	9,17	11,67	22,22	32,11	24,00

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20000219	28,06	23,06	1008,3	10,46	10,93	18,33	33,00	24,28
20000220	27,83	22,94	1008,5	9,98	14,26	27,78	32,22	24,72
20000221	27,72	23,72	1008,7	9,33	9,82	24,08	31,39	24,61
20000222	27,72	23,56	1009,1	7,24	5,37	18,33	32,61	24,61
20000223	28,00	23,67	1008,5	7,40	5,37	18,33	32,39	24,39
20000225	26,06	23,22	1007,4	6,28	8,33	18,33	31,61	23,00
20000226	26,28	23,50	1007,3	7,72	12,04	22,22	30,78	23,00
20000227	27,39	23,50	1007,0	7,89	7,78	22,22	32,22	24,00
20000228	27,61	23,72	1008,2	9,33	10,56	14,82	31,00	24,39
20000229	27,94	23,17	1007,3	8,69	13,89	18,33	31,00	25,22
20000301	28,33	23,39	1007,1	8,85	16,30	27,78	31,00	25,28
20000302	26,28	23,28	1007,6	8,37	11,48	33,52	30,00	25,00
20000308	27,28	23,22	1009,8	10,30	9,26	18,33	32,00	14,39
20000309	27,94	23,33	1008,9	9,17	10,37	18,33	31,61	24,00
20000310	27,72	23,61	1009,2	9,17	5,93	14,82	32,22	25,00
20000314	27,89	22,72	1007,7	9,66	7,04	24,08	33,22	22,39
20000315	28,28	22,06	1006,6	9,66	13,52	22,22	31,22	22,50
20000316	29,17	23,11	1005,3	10,14	19,63	27,78	33,00	25,00
20000317	28,50	23,17	1005,7	10,62	14,45	37,04	33,72	24,78
20000318	28,06	23,39	1007,2	8,05	10,19	27,78	33,89	24,39
20000319	27,72	23,72	1008,0	8,69	8,15	14,82	31,00	24,22
20000320	27,89	23,50	1008,3	8,85	4,07	14,82	31,50	24,00
20000321	28,61	23,56	1009,6	7,89	7,22	14,82	32,11	23,72
20000325	26,83	23,00	1009,0	8,85	8,70	18,33	30,78	24,00
20000327	25,89	23,06	1008,9	7,40	5,00	11,11	30,72	23,39
20000328	27,06	24,17	1009,7	10,30	2,96	11,11	31,22	23,11
20000329	27,94	23,89	1009,8	7,56	5,56	16,48	32,00	24,00
20000330	28,72	23,94	1007,5	7,40	4,44	14,82	32,00	25,78
20000401	27,78	24,44	1007,2	7,89	2,22	11,11	30,00	24,39
20000404	27,06	23,89	1007,8	6,60	4,44	11,11	32,00	24,61
20000405	29,00	23,83	1007,0	9,66	7,59	11,11	31,61	24,28
20000406	27,22	24,00	1007,7	7,72	9,07	18,33	30,39	23,78
20000408	28,39	24,44	1007,5	9,33	8,70	22,22	31,61	25,28
20000410	28,11	23,61	1009,0	8,21	9,82	22,22	32,00	23,39
20000411	27,50	24,06	1008,2	6,60	4,82	11,11	31,39	23,39
20000412	26,94	24,11	1006,8	6,60	5,74	27,78	31,39	24,00
20000413	27,83	24,28	1007,2	6,60	8,33	18,33	31,22	23,39
20000414	27,56	23,94	1007,7	8,85	7,96	14,82	30,61	24,72

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20000415	26,50	23,78	1008,4	8,69	5,56	14,82	31,00	24,39
20000416	27,50	24,06	1007,9	9,50	5,74	18,33	31,39	24,28
20000417	28,72	23,61	1008,8	9,50	11,67	25,93	32,61	24,72
20000418	28,33	23,67	1009,3	9,01	6,30	18,33	32,00	24,61
20000419	27,39	23,72	1008,4	8,37	8,33	27,78	32,78	24,00
20000420	28,44	23,56	1008,7	7,89	9,07	18,33	32,00	24,28
20000422	27,11	24,11	1009,1	9,01	6,85	14,82	30,22	23,00
20000423	28,89	24,56	1008,6	10,30	7,04	18,33	31,50	24,39
20000424	25,94	23,89	1009,8	8,21	5,19	29,45	28,00	24,61
20000425	27,39	24,11	1008,7	7,89	6,67	33,52	31,50	24,39
20000426	26,44	23,67	1009,1	7,08	4,63	11,11	31,22	24,78
20000427	28,56	23,83	1009,4	7,08	6,30	14,82	32,00	24,61
20000428	28,39	24,11	1009,6	8,37	5,74	18,33	32,00	24,22
20000429	29,17	22,83	1010,1	9,33	11,48	18,33	32,00	24,78
20000503	27,11	23,56	1010,9	8,69	10,74	18,33	30,28	23,78
20000504	28,22	24,06	1010,3	10,30	8,52	18,33	30,11	23,61
20000506	28,61	23,39	1011,3	9,01	9,63	24,08	32,78	24,78
20000507	28,28	22,28	1010,6	9,01	11,85	37,04	31,61	24,00
20000508	28,78	23,22	1010,2	7,72	6,85	18,33	32,00	24,28
20000509	28,50	22,89	1009,4	9,33	9,07	18,33	32,00	24,00
20000510	28,89	23,11	1008,4	10,62	9,45	18,33	32,22	23,00
20000511	29,33	23,17	1007,9	9,66	10,37	18,33	32,39	24,00
20000512	29,50	23,50	1008,1	9,82	9,45	18,33	32,00	23,78
20000514	29,44	23,78	1009,6	10,30	11,85	18,33	32,00	25,39
20000515	29,06	25,28	1009,9	9,98	14,26	27,78	32,00	25,78
20000516	28,61	24,67	1010,0	10,94	12,78	24,08	31,00	25,61
20000517	29,22	23,78	1009,8	9,01	10,93	20,56	32,28	25,00
20000520	27,11	24,39	1010,3	7,89	3,70	12,96	31,61	24,22
20000522	28,50	24,72	1010,6	10,46	10,00	22,22	31,00	24,61
20000523	28,44	24,61	1010,7	10,14	5,74	14,82	32,00	25,00
20000524	29,17	23,83	1008,8	9,17	8,70	16,48	32,00	25,00
20000525	28,89	23,72	1009,0	8,53	10,74	20,56	32,39	25,28
20000527	28,67	23,67	1008,6	9,01	8,33	24,08	32,61	25,00
20000529	29,56	24,06	1008,8	10,46	14,45	29,45	32,00	24,78
20000530	29,78	22,94	1008,7	10,62	14,45	25,93	32,00	25,39
20000531	28,11	21,50	1007,5	10,30	12,59	27,78	32,00	23,00
20000601	27,28	21,83	1008,5	9,01	10,00	16,48	31,22	22,61
20000602	29,06	22,83	1007,7	9,98	19,63	33,52	31,28	23,50

Timestamp	TEMP	DEWP	SLP	VISIB	WDSP	MXSPD	MAX	MIN
20000603	27,06	23,56	1008,4	8,53	12,22	18,33	31,00	24,00
20000604	27,50	24,50	1008,4	9,50	5,93	16,48	30,00	23,61
20000605	28,17	23,89	1009,3	9,17	10,74	22,22	31,22	24,78
20000607	28,11	22,61	1010,7	9,50	11,11	24,08	31,39	22,00
20000608	27,72	21,83	1011,4	9,82	8,52	18,33	32,00	23,61
20000609	28,94	21,00	1010,6	9,82	9,45	22,22	32,00	22,61
20000610	27,83	22,50	1011,4	8,21	7,59	14,82	31,00	23,61
20000611	27,61	22,06	1011,1	8,69	8,89	18,33	31,39	24,00
20000612	27,17	22,39	1011,2	7,40	6,85	16,48	31,39	23,00
20000613	28,00	22,78	1011,0	7,40	7,59	16,48	32,00	23,28
20000614	28,22	21,44	1011,4	9,98	8,33	14,82	31,61	23,72
20000615	27,67	21,11	1010,8	9,50	15,00	31,30	31,22	22,39
20000616	27,78	21,44	1010,7	8,69	9,07	18,33	31,00	22,22
20000618	29,17	22,72	1009,6	9,98	20,74	37,04	31,39	25,22
20000619	26,39	22,50	1011,1	9,82	14,82	27,78	28,39	25,00
20000620	28,22	22,72	1011,3	7,89	7,78	16,48	31,39	24,11
20000621	28,28	22,17	1011,7	9,66	11,85	18,33	31,39	24,00
20000622	27,72	22,83	1012,2	8,53	9,26	20,56	31,00	23,22
20000623	28,50	22,06	1010,8	9,98	14,63	31,30	31,39	22,78
20000624	28,17	20,33	1010,7	9,82	12,59	24,08	31,00	22,39
20000625	27,67	20,06	1009,9	9,33	10,00	22,22	31,72	21,39
20000626	26,33	20,11	1009,6	9,33	6,48	18,33	32,00	20,28
20000627	27,28	20,22	1009,5	10,14	7,04	18,33	31,28	21,28
20000628	27,44	20,83	1009,0	8,05	8,15	20,56	31,39	21,78
20000629	27,83	21,17	1008,2	7,24	9,26	16,48	31,22	22,78
20000630	26,89	20,89	1008,6	9,66	8,52	22,22	31,00	22,00
20000701	27,56	22,00	1008,5	9,17	12,22	25,93	30,11	22,11
20000702	27,33	21,94	1007,6	8,53	8,52	18,33	30,28	23,00
20000703	28,33	21,22	1007,3	9,33	10,19	18,33	31,22	23,78
20000704	28,39	22,00	1008,4	9,17	8,52	14,82	31,89	23,00
20000705	28,33	21,89	1009,2	9,33	9,45	14,82	31,61	23,00
20000706	28,33	21,22	1010,3	9,66	10,56	25,93	32,00	24,00
20000707	29,67	21,67	1010,6	9,50	9,26	18,33	31,00	24,00
20000708	28,06	22,00	1009,8	9,66	8,52	18,33	31,00	23,00
20000709	27,89	20,83	1009,8	9,01	12,59	18,33	31,00	22,50
20000710	27,17	19,44	1010,7	9,33	9,82	24,08	31,00	21,72
20000711	27,72	19,61	1010,0	8,53	7,96	18,33	31,39	21,39
20000712	28,44	21,72	1010,6	8,69	14,08	25,93	31,39	21,61

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20000713	28,78	21,61	1011,0	9,66	16,11	37,04	31,50	21,61
20000714	26,89	20,67	1010,7	9,82	15,74	33,52	30,89	21,00
20000715	27,89	19,39	1010,0	9,01	25,00	37,04	30,39	21,00
20000717	28,44	21,61	1011,8	10,14	27,59	37,04	31,00	24,50
20000718	28,67	21,50	1010,7	10,14	18,71	27,78	31,00	24,89
20000719	28,89	21,78	1010,1	9,33	16,85	33,52	31,00	24,00
20000720	27,83	21,78	1010,7	8,21	9,45	18,33	31,00	23,00
20000721	27,00	20,44	1011,5	8,69	17,22	35,93	30,50	22,50
20000722	27,22	17,94	1010,9	9,50	12,22	25,93	30,50	22,00
20000723	27,56	18,89	1010,4	9,33	16,11	27,78	31,00	21,39
20000724	27,28	20,83	1011,2	9,66	14,26	24,08	30,28	20,22
20000725	29,39	21,17	1010,4	9,50	12,96	22,22	32,00	21,39
20000726	27,72	21,94	1010,2	8,53	10,00	18,33	31,61	23,11
20000727	28,22	22,00	1010,0	9,98	15,37	27,78	31,39	24,22
20000729	27,94	22,50	1010,6	9,66	19,82	31,30	31,00	23,22
20000730	28,11	21,61	1010,4	9,17	13,15	27,78	31,00	23,00
20000731	28,89	21,61	1010,0	9,82	10,56	18,33	32,11	23,00
20000801	27,28	21,50	1010,2	9,50	11,30	33,52	31,61	22,00
20000802	26,94	20,50	1010,3	9,66	13,33	29,45	31,89	22,39
20000803	28,00	21,44	1010,3	8,69	12,78	29,45	31,39	22,00
20000804	27,67	20,39	1010,9	8,37	15,74	27,78	30,78	22,00
20000805	27,22	20,11	1010,9	9,98	13,33	31,30	31,39	21,78
20000806	27,39	21,28	1010,8	8,53	10,56	22,22	31,22	21,72
20000807	28,06	22,17	1010,9	9,33	12,59	27,78	32,00	21,28
20000808	27,56	20,28	1010,6	9,01	14,45	27,78	31,22	22,39
20000809	27,67	21,28	1010,1	9,50	11,48	27,78	31,28	22,00
20000810	27,83	20,78	1009,3	9,01	10,37	27,78	32,00	22,00
20000811	28,44	19,67	1010,0	9,33	13,70	25,93	31,22	21,39
20000812	28,94	19,39	1012,0	10,30	19,26	35,19	31,22	27,00
20000813	27,11	21,67	1011,2	9,01	12,78	27,78	30,50	22,00
20000814	28,22	19,17	1013,4	8,05	12,96	27,78	31,00	22,00
20000815	27,33	20,22	1011,0	8,21	9,45	22,22	31,61	22,50
20000816	27,61	20,33	1010,0	8,69	10,93	18,33	32,50	22,28
20000817	26,89	19,83	1010,0	9,33	9,07	22,22	31,00	22,39
20000818	28,33	20,28	1010,8	10,14	13,52	20,56	31,00	22,22
20000819	27,17	19,67	1010,5	9,50	9,07	18,33	31,39	22,72
20000820	27,44	19,50	1009,3	9,17	10,19	18,33	31,00	22,00
20000821	27,44	19,17	1008,2	8,53	12,96	29,45	31,00	21,00

A-6

Timestamp	TEMP	DEWP	SLP	VISIB	WDSP	MXSPD	MAX	MIN
20000822	27,56	17,44	1009,9	9,33	11,67	20,56	30,61	20,78
20000823	27,11	21,17	1010,2	9,82	12,96	27,78	30,39	20,72
20000824	29,33	21,89	1010,2	10,14	17,04	27,78	31,00	21,78
20000825	28,61	22,56	1010,0	9,50	14,45	22,22	31,22	23,39
20000826	29,50	21,17	1011,2	9,66	17,22	25,93	32,00	25,00
20000827	28,28	22,56	1011,3	9,82	12,96	27,78	31,50	24,00
20000828	29,44	21,06	1013,6	9,17	15,56	25,93	33,00	23,78
20000829	27,61	21,61	1009,7	9,50	10,56	20,56	32,00	24,00
20000830	30,56	19,83	1009,7	10,46	15,56	22,22	32,00	23,39
20000831	28,89	22,06	1010,6	9,82	20,00	27,78	31,22	25,00
20000901	29,17	21,39	1010,8	9,50	17,04	27,78	31,78	24,72
20000902	29,44	22,78	1010,8	9,33	18,15	27,78	33,00	25,22
20000903	28,83	21,06	1009,5	10,14	17,22	24,08	31,00	24,00
20000904	27,89	21,44	1010,0	8,05	7,59	18,33	33,00	23,61
20000905	29,94	21,78	1011,3	9,33	14,45	27,78	33,00	25,00
20000906	28,33	21,06	1010,7	9,33	11,48	27,78	32,00	22,00
20000907	29,61	21,11	1010,8	10,30	13,70	35,19	33,00	25,00
20000908	29,94	20,06	1011,1	9,98	14,26	27,78	33,28	22,61
20000909	27,78	20,33	1010,5	8,37	6,67	18,33	32,28	22,00
20000910	28,94	19,39	1009,7	9,50	12,78	22,22	33,00	22,11
20000911	28,72	20,17	1009,8	9,01	13,52	27,78	32,28	21,72
20000912	28,83	21,78	1010,3	8,85	13,52	24,08	32,22	22,72
20000913	28,67	22,11	1010,3	8,85	14,82	29,45	33,00	22,00
20000914	28,94	20,72	1009,8	9,17	21,11	31,30	32,00	22,28
20000915	28,94	19,83	1009,9	9,17	14,26	25,93	33,00	21,78
20000916	27,28	21,94	1010,2	9,01	8,52	25,93	32,61	21,61
20000917	29,78	21,28	1010,2	9,17	17,96	27,78	33,00	23,61
20000918	29,39	21,78	1008,9	8,37	12,78	27,78	32,28	25,00
20000919	29,78	22,56	1008,6	8,85	14,63	22,22	32,39	26,00
20000920	29,67	21,67	1009,3	8,69	11,67	22,22	33,00	25,00
20000921	28,67	22,61	1010,4	8,21	13,89	27,78	32,39	24,00
20000922	28,89	20,78	1010,8	9,33	16,48	27,78	33,00	24,00
20000923	29,56	21,61	1010,7	9,17	14,82	27,78	33,72	23,78
20000924	29,50	21,78	1010,4	9,50	13,15	25,93	33,61	25,00
20000925	30,00	21,11	1010,1	9,66	15,00	27,78	34,39	24,00
20000926	29,22	21,94	1010,9	8,85	8,33	27,78	33,39	24,28
20000927	29,44	21,22	1010,1	8,85	12,22	27,78	33,00	24,00
20000928	29,17	21,28	1010,5	9,66	12,96	27,78	34,00	24,00

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20000929	29,72	21,50	1011,1	8,85	15,19	27,78	33,00	24,00
20000930	28,83	21,44	1010,9	9,33	14,63	27,78	33,22	24,00
20001001	29,28	20,56	1010,5	9,33	12,59	27,78	34,00	24,22
20001002	30,06	22,22	1010,1	9,50	10,00	27,78	34,22	24,61
20001003	29,67	21,89	1008,7	9,01	10,37	18,33	34,00	25,00
20001004	29,67	22,11	1009,3	10,14	12,22	25,93	33,39	24,28
20001005	29,56	22,06	1009,5	8,37	11,67	27,78	34,28	24,00
20001006	29,89	21,78	1009,9	8,37	8,89	24,08	34,00	24,39
20001007	30,56	21,56	1009,3	9,98	11,11	22,22	34,00	24,00
20001008	29,44	21,72	1008,6	9,82	12,41	24,08	33,89	24,00
20001009	30,17	20,83	1008,0	9,01	11,67	24,08	34,00	23,39
20001010	29,56	22,11	1008,8	9,50	11,48	22,22	34,00	24,28
20001011	30,89	21,17	1009,2	9,66	15,00	24,08	34,39	25,78
20001012	30,22	22,17	1008,8	8,85	12,41	24,08	34,00	25,39
20001013	29,56	23,11	1006,9	8,85	10,19	24,08	33,89	25,00
20001014	29,50	23,61	1007,3	10,30	8,89	18,33	34,00	25,39
20001015	29,06	23,50	1009,1	9,33	6,85	18,33	33,78	25,61
20001016	28,50	23,28	1008,6	9,82	7,96	22,22	33,00	25,39
20001017	28,78	23,22	1007,3	8,05	11,30	22,22	32,11	24,78
20001018	28,50	23,50	1007,7	8,05	5,74	18,33	33,00	24,78
20001020	29,11	22,89	1012,2	9,98	10,74	22,22	33,00	24,00
20001021	28,94	22,33	1008,3	9,50	11,30	20,56	32,00	25,00
20001022	28,50	21,67	1012,3	10,94	13,33	18,33	33,00	24,00
20001024	27,89	22,56	1007,5	14,97	8,89	18,33	30,00	25,00
20001026	28,06	23,67	1009,5	9,50	8,52	20,56	31,00	25,00
20001028	26,78	24,11	1010,2	8,85	4,44	25,93	31,00	22,78
20001031	28,78	23,83	1008,3	9,33	9,07	18,33	31,78	24,78
20001102	28,39	24,17	1009,0	8,85	6,67	16,48	32,00	25,78
20001103	27,67	24,00	1010,4	8,53	7,04	27,78	33,00	25,00
20001104	29,28	23,61	1010,6	8,21	7,78	20,56	33,22	24,78
20001106	28,39	23,78	1008,1	9,66	7,22	22,22	32,78	24,22
20001107	29,39	23,61	1008,8	10,62	9,07	18,33	32,61	24,00
20001108	29,56	24,56	1008,2	9,66	10,74	37,04	32,78	25,39
20001110	29,44	24,44	1008,4	9,33	8,33	18,33	33,22	24,39
20001111	27,39	24,22	1008,3	7,24	7,41	16,48	31,39	24,61
20001112	29,33	23,94	1006,7	10,14	9,26	22,22	32,61	24,28
20001114	28,67	24,22	1006,1	9,17	9,07	31,30	32,39	24,28
20001115	28,06	24,61	1006,2	8,37	5,00	14,82	32,00	25,00

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20001117	29,11	24,11	1006,7	8,37	9,45	18,33	33,39	24,39
20001118	28,61	24,67	1006,1	8,53	4,44	14,82	32,22	25,50
20001119	27,78	24,17	1007,0	8,37	6,30	22,22	33,00	25,22
20001120	26,61	23,78	1007,2	6,28	8,89	27,78	32,39	23,00
20001123	27,94	23,17	1008,6	9,17	8,89	20,56	32,39	24,61
20001124	28,11	23,83	1008,3	7,56	5,93	27,78	31,00	24,72
20001125	29,17	24,50	1007,8	9,66	9,26	18,33	32,89	24,78
20001126	27,61	23,56	1008,9	9,66	10,37	25,93	32,78	24,00
20001127	26,11	23,67	1008,3	9,66	7,59	27,78	29,50	25,00
20001128	27,83	23,33	1008,0	9,82	13,52	18,33	31,50	24,39
20001129	28,11	23,89	1007,6	9,17	12,04	18,33	32,39	24,00
20001130	27,83	22,83	1007,5	9,98	13,33	24,08	30,78	24,28
20001201	30,11	22,72	1006,1	9,50	17,41	37,04	34,00	23,00
20001202	28,83	20,89	1007,1	9,82	11,67	18,33	34,22	24,61
20001203	28,00	20,67	1007,8	9,33	10,19	25,93	33,22	23,00
20001204	30,00	20,11	1008,0	9,98	11,67	16,48	34,39	23,22
20001205	28,28	20,61	1007,2	10,46	9,07	25,93	35,22	23,00
20001206	30,28	19,83	1013,7	9,98	16,11	37,04	34,00	23,22
20001207	30,78	21,06	1012,2	10,30	14,63	22,22	34,00	23,39
20001208	30,06	22,28	1006,3	9,82	15,56	37,04	35,22	24,11
20001209	28,39	23,39	1007,3	9,98	8,52	18,33	32,61	23,78
20001211	27,22	24,11	1008,4	9,17	5,56	18,33	32,39	24,72
20001212	28,11	24,00	1008,2	7,40	8,33	25,93	33,22	24,00
20001214	26,72	23,94	1008,8	6,92	5,19	14,82	29,22	24,50
20001215	27,11	24,06	1009,2	7,40	4,07	12,96	31,00	24,89
20001217	27,06	23,83	1008,5	7,08	10,00	27,78	31,39	24,00
20001218	27,83	24,06	1007,3	8,53	10,00	18,33	31,39	24,89
20001219	28,72	24,06	1007,1	6,92	9,07	20,56	31,39	24,78
20001221	27,72	24,39	1007,5	8,37	7,22	22,22	31,00	25,00
20001223	27,78	23,44	1009,6	9,98	5,56	18,33	31,00	25,00
20001224	28,67	23,00	1007,8	10,46	11,48	27,78	33,00	24,22
20001225	31,44	23,22	1008,6	9,82	10,93	18,33	35,00	24,78
20001226	30,06	22,17	1007,9	10,14	10,93	18,33	35,00	24,00
20001227	29,78	23,11	1008,0	10,46	12,22	22,22	32,78	25,00
20001228	29,50	23,67	1007,9	9,17	7,59	18,33	32,22	25,00
20001229	29,22	23,67	1007,5	9,98	6,67	12,96	34,22	25,00
20001230	30,17	23,72	1007,5	9,82	10,19	24,08	33,39	24,78
20001231	29,72	23,17	1007,7	10,62	11,67	18,33	34,00	25,00

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20010101	29.89	23.17	1008.2	9.82	11.11	18.33	33.50	25.00
20010102	28.33	23.33	1008.7	10.14	12.04	22.22	32.28	24.39
20010104	28.67	23.78	1008.9	9.98	6.11	14.82	32.61	25.11
20010105	26.33	23.72	1009.4	9.82	5.93	18.33	31.00	24.00
20010106	26.33	23.67	1009.6	7.40	6.30	22.22	31.72	23.72
20010108	24.89	23.17	1007.5	7.08	8.89	27.78	31.61	23.00
20010110	27.89	22.83	1007.4	9.82	20.74	35.19	32.00	23.78
20010111	26.06	22.39	1009.0	9.50	16.48	33.52	28.00	23.78
20010112	28.17	22.78	1011.2	10.14	7.96	22.22	30.00	24.00
20010113	28.78	22.72	1008.7	10.62	7.41	18.33	31.00	24.39
20010117	28.78	23.17	1007.6	9.01	17.22	24.08	33.61	24.11
20010119	27.00	24.11	1008.3	6.76	5.56	14.82	31.00	24.22
20010121	26.50	23.78	1008.4	6.92	9.07	16.48	31.39	22.89
20010123	27.39	23.61	1008.0	10.14	11.11	18.33	32.61	24.50
20010124	29.17	23.72	1007.9	9.01	7.41	18.33	31.72	24.61
20010126	28.06	24.94	1007.4	8.69	6.48	14.82	32.22	25.78
20010127	29.78	24.56	1009.5	10.62	8.15	18.33	33.00	26.00
20010128	27.06	23.67	1007.7	8.05	10.56	27.78	33.11	23.00
20010129	27.50	24.39	1007.3	7.40	14.26	22.22	32.78	24.39
20010130	27.56	24.11	1007.4	8.53	10.56	14.82	31.61	23.78
20010131	27.78	24.06	1006.5	8.21	5.37	18.33	31.78	23.28
20010204	26.17	22.39	1006.9	8.69	12.96	18.33	30.89	24.39
20010205	27.94	22.89	1005.5	9.33	15.19	27.78	32.22	24.28
20010206	26.50	23.61	1005.4	7.40	9.26	18.33	31.78	24.00
20010207	27.50	23.11	1004.5	8.85	15.56	27.78	33.00	24.00
20010208	26.17	22.83	1004.6	8.21	17.78	27.78	31.00	24.00
20010210	26.78	23.56	1003.5	7.56	18.52	27.78	31.61	23.78
20010212	28.06	22.72	1004.2	8.85	22.59	37.04	32.61	24.00
20010213	27.67	22.72	1004.5	8.69	19.26	37.04	31.72	24.00
20010214	28.94	22.94	1006.6	9.82	22.22	37.04	33.39	25.00
20010215	28.17	22.56	1007.6	9.01	18.89	33.52	33.22	24.89
20010216	29.39	22.11	1008.5	9.01	16.30	25.93	34.50	24.61
20010218	27.22	23.11	1008.6	8.37	11.11	31.30	33.50	23.61
20010219	27.28	22.56	1008.2	8.69	12.04	22.22	33.61	23.61
20010220	27.44	22.89	1007.4	7.72	11.48	27.78	32.39	24.00
20010221	28.22	22.17	1007.9	9.01	11.30	37.04	31.39	23.28
20010223	28.67	23.22	1009.1	9.98	10.19	14.82	31.39	25.89
20010225	26.56	23.72	1009.7	8.85	6.48	14.82	30.00	23.00

A-10

Timestamp	TEMP	DEWP	SLP	VISIB	WDSP	MXSPD	MAX	MIN
20010226	28,44	23,28	1010,1	9,17	8,52	14,82	31,78	23,89
20010227	27,44	24,17	1010,6	6,44	7,78	29,45	31,89	23,78
20010301	27,78	23,94	1010,1	6,76	8,70	20,56	32,11	24,00
20010303	27,33	24,06	1008,9	7,72	11,30	16,48	32,00	24,00
20010304	28,56	23,39	1007,8	10,30	11,30	29,45	33,78	24,00
20010305	25,33	23,39	1008,3	7,40	10,93	25,93	32,22	22,61
20010306	27,50	23,50	1008,8	8,37	9,07	14,82	32,22	22,61
20010307	27,11	23,50	1009,1	9,17	5,00	14,82	31,39	23,11
20010308	27,67	23,83	1008,9	7,56	6,85	18,33	31,28	24,28
20010309	27,28	23,67	1009,2	9,98	3,70	12,96	31,22	24,39
20010311	25,94	23,67	1010,0	6,92	6,30	14,82	32,39	23,00
20010312	26,56	23,33	1009,7	7,24	4,44	16,48	31,22	23,39
20010313	27,94	23,67	1009,9	8,69	5,93	12,96	32,28	23,22
20010314	28,78	23,17	1010,8	7,40	9,07	29,45	33,11	24,22
20010316	27,72	23,56	1008,7	8,69	7,22	14,82	30,72	24,00
20010317	27,44	24,50	1007,7	7,56	5,74	12,96	32,11	24,22
20010318	27,22	23,33	1007,9	8,21	9,63	18,33	32,61	24,11
20010319	26,78	23,22	1008,0	7,89	12,59	25,93	33,00	23,00
20010320	28,78	23,11	1007,0	10,14	15,56	22,22	33,00	23,50
20010321	28,33	22,94	1006,7	9,17	16,48	27,78	33,50	23,61
20010322	28,89	23,00	1007,0	9,82	7,04	12,96	32,11	24,22
20010324	26,89	24,39	1008,0	6,92	4,07	14,82	32,61	23,72
20010325	26,11	23,72	1008,2	7,56	4,82	16,48	32,00	24,00
20010326	27,44	24,00	1007,9	7,08	5,56	22,22	31,89	24,28
20010327	27,83	23,67	1008,8	8,69	7,96	18,33	31,39	24,00
20010328	27,56	24,28	1008,5	9,17	2,96	9,45	31,78	23,22
20010329	27,56	24,17	1008,0	8,69	5,56	14,82	32,11	24,22
20010330	26,83	24,56	1008,7	7,08	3,89	14,82	32,00	24,78
20010331	26,72	24,11	1010,2	6,44	3,52	12,96	31,00	24,00
20010402	27,83	24,33	1008,7	9,66	7,22	18,33	31,00	23,50
20010403	28,33	24,50	1008,2	9,66	7,59	14,82	31,78	24,22
20010404	27,28	24,17	1009,3	8,53	7,59	14,82	31,61	24,00
20010405	28,06	24,11	1009,2	9,17	5,74	18,33	31,28	24,00
20010406	28,67	24,61	1010,1	9,66	3,33	14,82	32,72	24,11
20010407	28,50	24,83	1009,7	8,37	5,56	12,96	32,61	24,78
20010408	27,89	24,50	1008,5	8,85	6,30	22,22	31,89	24,39
20010412	27,28	23,94	1009,4	7,72	6,11	18,33	32,61	24,00
20010414	27,22	24,06	1011,6	7,24	6,30	14,82	31,39	24,00

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20010415	27,67	24,17	1011,1	9,50	11,11	20,56	29,00	24,00
20010416	27,00	24,06	1009,5	8,85	9,07	24,08	31,28	23,78
20010417	27,00	23,67	1009,4	8,05	6,30	12,96	30,61	23,61
20010418	27,17	22,72	1008,4	8,37	7,04	14,82	31,50	23,00
20010419	29,06	21,94	1008,2	9,50	6,30	11,11	31,72	23,39
20010420	28,22	22,11	1007,8	9,01	6,48	11,11	31,61	24,00
20010421	29,00	22,28	1007,7	9,33	10,56	20,56	31,72	24,22
20010422	29,28	22,33	1007,6	9,01	7,22	12,96	32,61	24,22
20010425	28,78	25,00	1008,4	9,17	10,19	14,82	31,39	24,39
20010426	29,28	24,06	1008,2	9,01	9,63	14,82	32,72	25,00
20010427	28,33	23,33	1008,0	8,21	5,19	12,96	32,39	23,00
20010428	28,72	22,89	1008,1	8,69	5,93	14,82	32,22	23,39
20010429	29,89	21,11	1008,9	9,50	7,22	11,11	32,72	24,39
20010501	28,50	24,61	1009,5	8,21	7,78	14,82	32,61	24,78
20010502	28,06	24,44	1008,7	8,21	5,74	12,96	32,39	24,39
20010504	29,39	23,94	1010,4	10,14	9,45	14,82	32,39	25,00
20010505	29,78	23,83	1010,3	7,72	6,30	12,96	33,00	24,89
20010506	29,00	22,83	1009,8	9,17	10,74	22,22	32,39	24,78
20010508	29,61	22,56	1009,3	10,46	9,07	14,82	32,89	25,11
20010510	28,17	21,61	1010,6	10,46	10,37	18,33	31,39	23,00
20010511	27,83	20,83	1010,2	9,66	7,41	18,33	31,61	22,00
20010512	27,78	21,28	1009,2	13,04	5,93	11,11	31,72	22,22
20010513	29,11	21,50	1009,2	9,50	10,00	14,82	31,61	23,00
20010514	28,83	22,17	1010,4	9,50	7,41	12,96	31,89	24,28
20010515	28,72	22,56	1010,9	10,94	12,04	18,33	32,22	20,00
20010516	28,89	23,06	1010,9	10,14	12,22	22,22	31,78	24,22
20010517	29,56	23,61	1010,9	9,66	10,19	18,33	32,72	25,22
20010518	28,72	24,28	1011,3	9,98	11,30	25,93	33,00	25,00
20010519	27,78	24,11	1011,1	8,21	3,52	9,45	33,00	25,00
20010520	28,28	24,00	1009,4	9,01	6,48	11,11	31,00	25,28
20010521	29,22	23,28	1009,4	12,23	10,37	14,82	32,00	25,22
20010522	29,83	22,61	1008,9	13,36	11,67	14,82	32,00	24,00
20010523	30,00	22,28	1009,7	14,81	9,07	11,11	32,00	25,00
20010524	28,00	22,17	1008,9	9,66	7,04	12,96	32,22	23,00
20010525	28,94	21,56	1009,9	14,00	12,96	14,82	32,00	23,00
20010526	29,06	23,22	1007,9	13,52	14,45	22,22	32,00	24,00
20010527	29,44	24,17	1008,3	10,94	11,67	14,82	32,00	24,00
20010529	29,00	23,50	1008,1	33,80	12,96	14,82	31,00	25,00

A-12

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20010530	29,61	23,89	1009,9	9,98	9,07	9,45	31,00	27,00
20010531	29,44	23,83	1010,9	11,10	11,48	16,48	32,00	25,00
20010602	29,33	24,33	1011,1	12,39	10,56	12,96	31,00	25,00
20010603	29,61	23,56	1009,0	13,52	11,85	18,33	32,00	26,00
20010604	28,28	23,83	1012,6	11,27	12,59	14,82	31,00	24,00
20010605	27,89	23,28	1007,2	14,00	11,11	24,08	31,00	24,00
20010606	28,56	23,61	1007,9	13,36	9,26	12,96	31,00	25,00
20010607	29,17	23,56	1007,6	9,98	8,33	14,82	32,00	25,00
20010608	28,28	23,11	1007,5	11,91	9,63	12,96	32,00	25,00
20010611	27,28	24,11	1009,7	9,82	7,59	12,96	31,22	23,00
20010612	27,06	23,61	1010,6	15,13	9,45	14,82	30,00	24,00
20010613	28,44	24,11	1010,5	10,46	7,59	11,11	31,28	24,78
20010614	29,28	23,83	1010,0	8,69	5,37	11,11	32,28	25,22
20010616	28,78	22,50	1009,7	10,78	11,48	18,33	32,00	24,00
20010617	27,28	21,56	1010,4	9,66	11,48	18,33	32,00	23,00
20010618	28,00	19,89	1009,8	9,33	6,67	11,11	32,00	22,22
20010619	27,61	19,83	1009,6	12,23	10,74	18,33	31,00	22,11
20010620	27,67	21,06	1009,5	9,50	11,11	18,33	32,00	22,22
20010623	27,78	23,78	1009,7	13,04	12,78	18,33	30,00	24,22
20010624	29,11	23,22	1010,3	11,91	13,52	18,33	31,00	25,00
20010625	27,94	22,61	1011,3	10,46	7,41	12,96	30,00	24,39
20010626	26,94	21,39	1009,0	8,85	7,78	12,96	31,39	21,72
20010627	28,00	21,67	1010,6	11,91	9,82	18,33	31,22	22,78
20010628	27,56	21,94	1011,3	13,20	6,48	11,11	31,78	23,00
20010629	27,83	22,00	1011,1	9,33	7,78	12,96	31,22	23,00
20010630	27,89	22,22	1011,5	10,30	10,37	14,82	31,78	23,00
20010701	27,17	17,61	1011,1	9,33	11,30	27,78	32,00	20,78
20010702	26,11	19,11	1010,9	9,50	10,19	24,08	31,00	20,28
20010703	25,33	17,28	1012,7	9,98	10,74	18,33	30,00	19,00
20010704	25,94	17,11	1013,4	9,98	16,67	24,08	29,00	19,00
20010705	27,22	17,00	1013,7	12,39	10,93	22,22	30,00	20,00
20010706	27,61	19,83	1012,8	10,78	8,70	14,82	31,00	22,00
20010707	27,39	21,83	1008,7	11,10	11,67	18,33	30,00	23,00
20010708	27,22	22,11	1010,9	18,02	9,07	14,82	30,00	24,00
20010709	28,22	22,44	1010,8	17,54	11,67	14,82	30,00	24,00
20010710	28,61	23,28	1014,2	10,14	12,22	14,82	30,00	24,00
20010714	28,06	23,17	1011,2	9,98	16,11	22,22	30,78	23,89
20010715	28,44	23,00	1012,2	9,50	14,45	35,93	30,72	25,22

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20010716	27,56	23,17	1011,5	9,98	10,56	18,33	30,78	22,89
20010717	28,00	23,00	1010,8	10,46	8,52	14,82	31,11	24,00
20010718	27,50	22,28	1010,6	9,17	9,63	14,82	31,39	24,11
20010720	26,94	22,83	1009,5	8,69	8,15	14,82	29,78	23,78
20010721	27,78	23,17	1008,9	11,59	13,33	18,33	31,00	24,00
20010722	28,06	23,28	1009,5	11,59	11,30	16,48	30,78	25,61
20010723	27,61	22,67	1009,2	11,27	17,59	27,78	30,89	25,00
20010724	28,39	22,22	1008,7	9,98	10,19	14,82	31,61	24,00
20010725	28,22	21,39	1009,6	9,17	13,33	20,56	31,22	23,61
20010726	28,22	21,28	1010,2	9,50	13,15	25,19	31,28	23,22
20010727	28,11	20,72	1010,8	9,66	9,82	14,82	31,00	22,00
20010728	27,78	22,39	1010,8	9,17	12,04	25,93	31,78	22,28
20010729	27,39	21,72	1011,7	9,66	11,30	18,33	31,00	22,00
20010730	27,39	20,56	1011,8	9,01	8,89	14,82	31,39	22,00
20010731	26,94	20,00	1011,1	8,21	9,82	18,33	30,00	21,00
20010801	26,94	19,67	1009,6	9,33	8,15	16,48	30,50	20,39
20010802	26,78	19,89	1009,7	8,37	7,04	11,11	32,28	20,78
20010803	25,17	19,06	1011,0	7,89	7,22	12,96	30,00	20,00
20010804	25,61	18,83	1012,9	8,05	11,30	22,22	30,11	20,00
20010805	26,17	18,33	1012,6	8,53	10,19	14,82	30,11	19,78
20010806	27,56	18,39	1011,6	8,69	10,37	14,82	31,22	20,61
20010807	26,83	20,72	1011,2	8,85	11,30	20,56	30,22	21,50
20010808	27,33	21,50	1011,8	9,17	11,11	18,33	31,78	21,78
20010809	27,44	21,67	1012,6	9,01	10,74	16,48	31,39	22,00
20010810	28,22	22,11	1011,6	9,17	9,82	16,48	32,11	23,00
20010811	27,11	21,11	1012,0	8,69	7,96	14,82	32,00	22,00
20010812	27,28	20,61	1011,5	9,01	10,93	18,33	31,39	22,00
20010813	27,33	19,89	1012,5	8,37	9,45	14,82	30,61	22,28
20010814	27,72	20,33	1013,6	9,01	10,37	14,82	31,50	21,78
20010815	27,44	21,83	1012,8	8,85	11,30	22,22	31,39	23,00
20010816	27,56	21,00	1010,5	9,98	9,45	18,33	32,00	22,00
20010817	26,94	18,44	1010,6	9,82	7,59	14,82	31,00	21,61
20010818	26,94	18,44	1010,8	9,50	8,33	12,96	30,78	20,00
20010819	27,28	20,44	1010,9	9,01	11,48	20,56	30,39	21,61
20010820	28,22	20,72	1011,5	9,82	10,93	14,82	31,00	22,61
20010821	26,61	20,11	1011,1	8,85	10,37	22,22	31,22	22,00
20010822	27,94	20,50	1010,6	9,33	9,63	14,82	31,28	21,72
20010823	27,89	20,67	1009,9	9,17	11,48	16,48	31,72	21,61

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20010824	26,33	20,11	1010,5	8,69	11,85	22,22	31,39	22,00
20010825	27,56	19,94	1010,9	8,69	10,37	16,48	31,39	22,00
20010826	27,28	21,00	1010,5	9,01	10,37	18,33	31,61	22,00
20010827	26,11	20,22	1009,2	8,37	10,19	28,71	31,22	21,00
20010828	27,56	19,00	1009,4	8,85	10,56	14,82	31,78	21,72
20010829	27,78	21,67	1010,0	9,33	14,63	22,22	31,00	21,22
20010830	28,11	23,06	1010,0	9,50	17,04	27,78	31,11	23,22
20010831	28,72	23,00	1009,9	8,53	10,19	18,33	32,00	23,72
20010901	28,44	21,39	1009,6	8,85	10,93	16,48	32,22	23,39
20010902	27,78	22,17	1010,5	9,33	12,59	20,56	32,00	25,00
20010903	28,17	22,22	1010,8	9,50	10,37	18,33	31,61	24,72
20010904	28,56	22,39	1010,8	10,14	11,11	14,82	31,22	24,39
20010905	27,61	22,06	1011,3	8,21	9,45	14,82	31,00	23,00
20010906	28,67	19,94	1012,1	8,69	10,00	18,33	33,39	23,00
20010907	28,78	21,61	1012,3	9,50	11,30	18,33	32,78	23,00
20010908	27,78	21,17	1008,6	8,85	11,67	16,48	32,00	22,00
20010909	28,28	21,83	1011,6	8,53	10,19	12,96	32,00	23,00
20010910	27,83	20,61	1010,0	9,17	9,82	14,82	32,39	23,00
20010911	27,56	21,00	1011,1	9,01	10,37	20,56	33,00	22,00
20010912	28,61	21,39	1011,5	9,17	7,59	12,96	33,00	22,39
20010913	28,56	21,06	1010,1	9,50	10,19	22,22	32,78	23,00
20010914	28,06	21,50	1008,8	8,37	12,59	18,33	33,00	23,00
20010915	30,17	20,39	1010,3	9,66	11,48	20,56	33,00	24,00
20010916	30,33	20,78	1008,6	8,69	15,74	25,93	33,00	25,00
20010917	28,94	22,56	1010,0	9,01	12,41	18,33	33,61	24,00
20010918	29,39	22,78	1010,5	9,01	9,82	14,82	33,78	24,11
20010919	29,00	21,56	1009,9	9,50	9,63	14,82	33,89	24,00
20010920	27,11	22,00	1010,7	8,53	11,30	18,33	32,61	23,00
20010921	29,17	20,89	1010,1	9,33	11,48	18,33	33,28	23,72
20010922	29,00	22,33	1010,0	9,50	10,56	18,33	33,22	23,61
20010923	29,94	21,17	1010,4	9,66	11,30	25,93	34,00	24,00
20010924	29,22	21,06	1009,8	8,85	11,85	20,56	33,78	24,00
20010925	28,50	22,33	1009,6	8,85	11,11	18,33	33,00	23,39
20010926	29,33	22,00	1008,9	9,82	15,19	31,30	33,22	24,28
20010928	29,06	21,39	1009,7	9,17	10,00	18,33	33,50	24,00
20010929	28,89	21,67	1009,7	8,53	10,37	16,48	33,00	24,22
20010930	29,50	21,00	1010,1	9,33	10,19	14,82	33,61	25,00
20011001	29,17	22,22	1010,1	8,85	6,67	11,11	33,39	24,00

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20011002	28,72	22,72	1010,3	7,56	7,59	14,82	33,00	24,50
20011003	28,89	22,39	1009,5	7,89	10,37	18,33	34,61	24,00
20011004	28,83	22,22	1010,6	8,05	6,30	14,82	33,61	24,00
20011005	29,11	22,06	1009,3	9,33	11,11	14,82	34,22	24,78
20011006	30,22	20,39	1009,5	8,53	9,07	12,96	33,22	24,50
20011007	29,17	21,83	1010,9	8,21	9,45	18,33	34,11	24,39
20011009	26,83	22,89	1013,6	4,67	6,11	9,45	29,11	24,00
20011010	27,33	23,28	1009,3	8,37	8,89	12,96	29,00	25,00
20011011	28,56	22,00	1010,0	10,30	10,93	14,82	32,00	24,00
20011012	29,06	20,94	1009,6	10,78	10,19	16,48	32,00	24,22
20011013	27,00	22,28	1014,3	8,21	7,04	9,45	29,78	23,89
20011015	29,11	23,83	1010,6	10,14	9,26	12,96	31,00	23,61
20011016	29,78	23,33	1009,0	9,82	9,45	14,82	33,00	24,72
20011018	29,72	23,33	1010,2	8,21	10,00	18,33	34,00	25,00
20011019	29,61	23,11	1010,4	8,37	8,33	12,96	33,00	25,00
20011020	28,06	22,83	1009,8	7,89	10,00	14,82	33,00	25,00
20011025	27,67	24,56	1007,9	7,89	8,52	12,96	31,00	24,22
20011026	29,78	24,28	1007,6	10,46	9,07	12,96	33,00	25,22
20011027	30,78	23,44	1011,1	9,33	10,74	18,33	33,00	25,22
20011028	30,44	22,50	1012,6	9,66	9,07	11,11	33,28	25,00
20011029	30,44	22,44	1009,2	9,33	9,07	12,96	33,11	23,61
20011030	29,50	21,67	1010,2	8,85	9,07	14,82	33,72	24,00
20011031	27,94	22,56	1011,4	8,85	10,00	22,22	32,00	24,11
20011101	28,17	23,28	1011,7	7,72	8,52	12,96	32,00	23,89
20011102	29,50	23,33	1010,9	8,21	8,33	11,11	33,39	24,72
20011103	29,11	23,67	1009,6	7,24	6,48	12,96	33,89	25,00
20011104	30,28	23,00	1009,5	8,85	10,56	22,22	33,78	25,28
20011105	29,94	23,67	1010,1	8,21	6,67	14,82	33,78	25,00
20011106	29,89	22,78	1009,0	9,33	6,48	12,96	34,00	24,78
20011107	28,72	23,78	1008,2	6,12	5,19	12,96	33,22	24,89
20011108	30,50	23,83	1008,2	9,33	7,78	18,33	33,00	25,00
20011109	30,11	23,72	1008,9	9,01	9,45	14,82	33,00	25,72
20011110	30,28	22,94	1008,5	9,98	13,70	31,30	33,61	25,61
20011111	29,39	23,33	1008,5	9,50	5,37	9,45	33,22	25,39
20011112	30,94	22,56	1009,1	10,14	10,74	29,45	33,28	25,39
20011113	29,56	23,22	1009,4	9,01	9,63	16,48	33,22	25,39
20011114	28,00	24,28	1009,8	7,89	7,22	12,96	33,50	25,00
20011115	26,83	24,11	1009,8	9,33	7,78	18,33	31,39	24,61

A-16

Timestamp	TEMP	DEWP	SLP	VISIB	WDSP	MXSPD	MAX	MIN
20011116	27,17	23,44	1010,3	7,89	8,52	14,82	32,39	24,00
20011118	27,50	24,50	1009,5	7,08	4,26	11,11	30,78	23,28
20011119	27,11	24,17	1009,2	8,21	5,37	14,82	31,28	24,39
20011120	28,89	24,28	1009,3	10,14	9,07	12,96	32,11	25,28
20011121	29,83	24,39	1009,8	9,66	12,04	14,82	32,61	26,00
20011122	28,89	24,11	1009,7	9,01	5,93	14,82	33,22	25,00
20011123	29,06	24,00	1008,5	7,40	5,74	11,11	33,78	25,00
20011124	27,50	23,56	1007,6	8,05	11,30	31,30	34,00	24,00
20011125	27,83	23,50	1008,4	7,24	4,07	9,45	32,22	24,72
20011126	28,28	23,44	1009,5	7,40	4,63	11,11	32,39	24,72
20011127	29,28	23,39	1010,0	8,21	10,56	18,33	32,22	24,28
20011128	27,11	23,39	1009,3	8,21	5,74	11,11	33,11	25,00
20011202	26,89	22,83	1007,5	9,01	5,56	9,45	30,78	23,50
20011203	28,00	22,83	1007,3	9,33	6,85	12,96	32,00	23,78
20011204	28,72	23,00	1007,3	8,85	6,67	12,96	33,00	24,00
20011205	29,28	22,11	1007,1	10,14	8,52	27,78	33,11	23,78
20011206	28,33	22,44	1007,2	9,17	6,67	12,96	33,00	24,00
20011207	27,89	21,78	1011,3	8,53	9,45	14,82	33,00	24,00
20011208	26,72	21,39	1007,9	8,05	5,00	9,45	32,39	24,00
20011209	28,22	21,89	1008,9	8,69	7,04	12,96	32,00	24,00
20011211	28,11	22,67	1009,1	7,89	8,33	14,82	33,00	24,61
20011214	27,22	23,89	1011,2	8,37	5,19	12,96	31,22	25,00
20011216	26,50	24,17	1010,2	8,37	5,93	9,45	31,00	23,89
20011217	26,11	23,56	1010,6	7,89	6,11	12,96	32,00	23,00
20011218	28,33	23,50	1011,2	9,17	7,96	14,82	33,22	23,61
20011219	26,61	23,83	1010,3	7,24	7,96	16,48	32,00	24,00
20011221	28,11	23,61	1010,7	9,66	7,78	12,96	31,39	24,11
20011223	28,11	23,28	1011,3	9,33	6,30	12,96	32,61	24,22
20011224	28,78	23,61	1011,3	8,37	7,59	12,96	32,72	25,00
20011225	28,33	23,83	1011,0	8,37	8,33	17,96	32,50	24,78
20011226	28,22	23,56	1012,6	8,05	5,37	11,11	32,22	25,00
20011227	27,50	23,44	1012,8	7,72	5,37	11,11	32,50	25,00
20011228	27,28	24,00	1012,6	7,56	7,59	14,82	31,72	24,61
20011230	27,78	23,56	1010,4	8,21	7,96	12,96	33,39	23,78
20020102	27,17	23,61	1011,7	8,37	8,15	14,82	32,39	23,61
20020103	27,83	23,28	1011,2	8,53	11,30	18,33	32,78	23,00
20020104	28,44	23,83	1010,4	9,17	9,63	14,82	32,22	23,00
20020107	27,78	23,39	1011,2	8,05	6,30	11,11	31,89	24,00

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20020109	28,00	24,22	1010,8	6,76	18,33	30,40	31,61	24,72
20020110	28,44	23,33	1010,6	8,21	8,52	14,82	32,39	24,00
20020111	27,67	23,28	1009,0	9,66	8,52	14,82	32,39	24,28
20020112	29,61	23,61	1008,1	9,50	8,33	12,96	32,78	25,00
20020113	28,00	23,94	1008,0	7,89	4,26	9,45	33,22	25,00
20020114	27,44	23,89	1008,7	7,72	6,67	12,96	31,78	24,00
20020115	28,17	24,44	1008,4	7,08	6,11	9,45	32,11	24,00
20020116	27,33	24,28	1008,3	5,95	6,67	12,96	32,39	24,39
20020118	27,78	24,83	1008,3	7,56	8,52	17,96	30,39	24,61
20020119	27,72	24,22	1008,9	9,17	7,04	22,22	31,39	25,00
20020121	26,61	23,61	1008,6	8,69	8,89	11,11	31,00	24,00
20020122	27,06	23,22	1009,2	7,72	7,96	12,96	32,78	23,89
20020123	27,33	23,67	1009,3	7,40	7,96	12,96	32,61	20,78
20020125	26,22	24,56	1008,8	5,95	4,26	11,11	31,78	23,89
20020126	25,22	23,67	1008,9	7,08	6,30	12,96	30,11	24,00
20020127	27,94	24,17	1008,9	9,01	8,89	17,96	32,39	25,00
20020128	26,50	23,83	1009,2	7,72	7,59	14,82	31,00	24,00
20020129	26,50	23,61	1010,5	7,24	12,22	25,93	32,39	23,61
20020130	24,72	22,67	1011,5	7,56	11,67	16,48	29,28	23,72
20020201	27,78	23,89	1010,9	7,40	7,59	12,96	36,22	24,22
20020202	27,22	23,56	1010,0	8,85	12,59	24,08	31,39	25,00
20020203	26,94	23,78	1008,6	8,37	15,74	27,78	31,61	23,61
20020204	27,06	23,61	1009,1	8,37	15,00	24,08	30,78	23,61
20020206	26,78	23,17	1009,2	9,82	16,48	29,45	32,11	24,00
20020207	27,89	22,17	1009,5	9,82	20,19	31,30	32,61	24,00
20020210	25,39	23,39	1009,1	5,15	13,70	31,30	29,28	24,28
20020211	28,28	23,67	1009,8	8,37	18,33	29,45	30,78	24,61
20020212	28,33	22,89	1009,8	8,53	19,08	25,93	33,39	24,00
20020213	26,50	23,00	1010,0	7,24	15,00	31,30	32,00	24,00
20020214	26,56	23,00	1010,5	7,56	16,30	22,22	28,39	23,39
20020215	27,28	23,11	1009,0	8,21	16,11	29,45	33,61	24,00
20020217	27,78	23,44	1013,3	9,66	21,30	31,30	30,00	23,00
20020218	27,83	23,39	1012,4	8,69	10,93	25,93	33,00	23,39
20020219	28,72	22,94	1011,9	9,01	16,48	25,93	33,39	24,39
20020220	26,94	23,50	1011,1	7,72	16,67	24,08	33,11	24,00
20020222	26,44	22,83	1014,1	8,05	20,19	37,04	31,28	24,39
20020224	27,94	23,83	1011,0	9,17	9,63	18,33	32,22	23,78
20020226	28,06	24,11	1010,5	7,89	6,48	18,33	31,61	23,61

Timestamp	TEMP	DEWP	SLP	VISIB	WDSP	MXSPD	MAX	MIN
20020228	27,94	24,11	1009,8	8,53	7,04	18,33	32,50	24,39
20020301	28,50	24,28	1008,9	4,67	6,30	18,33	32,00	24,78
20020302	26,83	23,67	1009,0	7,24	6,48	25,93	32,00	24,22
20020304	29,56	23,50	1008,2	8,85	8,70	14,82	32,39	24,78
20020305	27,94	23,83	1008,8	7,56	7,96	27,78	32,61	25,00
20020309	27,11	23,78	1010,9	6,92	8,52	18,33	31,89	24,00
20020311	26,83	23,78	1010,9	6,60	7,22	22,22	32,00	23,78
20020312	28,33	24,28	1011,7	8,53	7,41	22,22	32,00	23,72
20020313	29,67	23,67	1011,9	9,82	11,11	18,33	33,00	24,61
20020316	28,78	23,94	1010,5	8,85	8,33	14,82	32,00	25,00
20020318	28,17	24,44	1009,9	9,66	6,48	18,33	31,61	24,39
20020319	28,33	24,56	1010,2	9,66	7,22	22,22	31,00	24,61
20020320	27,72	24,22	1009,7	8,69	6,48	14,82	31,39	24,00
20020321	29,89	22,56	1008,4	9,66	8,15	14,82	32,00	24,39
20020322	29,28	24,56	1008,1	6,92	4,44	12,96	33,00	25,00
20020323	29,89	24,44	1007,2	7,56	5,56	18,33	33,00	24,78
20020324	29,89	24,00	1010,2	9,98	10,74	22,22	33,00	25,00
20020325	28,28	24,50	1007,4	8,05	4,82	18,33	33,00	25,39
20020326	29,56	24,67	1008,0	8,37	12,22	33,52	32,00	23,39
20020327	26,83	23,94	1009,2	7,72	10,00	37,04	32,00	24,00
20020328	27,61	24,28	1008,6	8,53	5,74	14,82	32,78	24,22
20020329	27,78	24,67	1010,1	8,21	5,93	14,82	31,72	24,78
20020330	27,06	24,50	1009,2	8,21	6,85	14,82	31,72	25,00
20020331	27,56	24,67	1009,7	7,56	8,15	14,82	30,00	25,39
20020401	29,00	23,17	1011,2	10,30	13,89	24,08	32,39	26,39
20020403	27,17	23,39	1008,8	7,89	7,59	18,33	32,39	24,00
20020405	29,33	24,50	1010,1	9,01	15,00	22,22	32,00	24,39
20020408	30,06	22,83	1010,3	9,33	12,96	18,33	33,00	25,00
20020409	27,56	23,11	1009,5	8,21	11,48	22,22	33,00	24,00
20020411	28,89	24,56	1007,1	6,60	4,26	11,11	33,39	26,00
20020412	27,50	24,33	1008,5	3,70	3,33	12,96	32,00	24,00
20020413	28,56	24,06	1009,5	8,69	4,82	12,96	31,11	25,00
20020414	28,94	24,06	1009,5	8,37	7,22	14,82	32,00	24,89
20020415	28,67	23,83	1010,0	7,56	9,07	18,33	32,22	24,78
20020418	28,00	24,06	1009,2	7,72	7,96	18,33	31,39	24,72
20020419	27,89	24,11	1010,0	9,82	12,22	22,22	31,00	24,00
20020420	27,00	23,78	1009,0	9,17	11,48	33,52	31,61	24,00
20020422	28,28	24,22	1008,2	8,69	9,26	14,82	31,78	24,61

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20020423	29,67	23,72	1008,8	8,85	10,74	18,33	31,78	25,00
20020424	29,06	24,06	1007,9	7,56	8,33	14,82	32,50	25,00
20020425	29,61	23,50	1007,5	9,50	11,85	27,78	33,00	25,11
20020426	29,28	23,67	1008,3	9,17	10,56	18,33	33,00	25,00
20020427	29,44	22,94	1008,5	9,01	13,52	24,08	32,39	25,39
20020428	29,83	23,33	1008,4	8,85	10,56	22,22	32,61	25,00
20020429	29,94	22,61	1007,2	9,66	11,85	18,33	32,28	24,78
20020430	29,28	22,83	1006,6	8,53	8,89	14,82	33,22	25,00
20020501	29,61	23,22	1007,1	7,89	8,70	14,82	33,00	25,00
20020502	30,06	24,11	1007,5	7,89	13,33	18,33	32,72	25,00
20020503	29,28	24,11	1007,1	9,50	20,19	33,52	32,78	24,61
20020504	30,17	24,33	1006,6	9,66	12,59	18,33	32,61	25,89
20020505	29,89	24,44	1006,4	9,01	12,59	22,22	32,61	25,72
20020506	28,94	24,22	1006,4	8,69	11,67	18,33	33,00	25,61
20020507	29,61	24,33	1007,6	8,53	11,67	18,33	32,72	25,39
20020508	29,33	24,00	1008,7	8,69	8,52	14,82	33,00	26,00
20020510	27,06	24,00	1010,2	5,95	7,59	14,82	31,72	24,00
20020512	29,44	24,89	1010,5	9,98	9,82	18,33	32,11	25,11
20020513	28,89	24,67	1011,0	9,33	7,59	25,19	32,00	25,11
20020514	29,00	23,72	1011,4	8,69	10,37	18,33	31,78	25,28
20160101	30,00	24,50	1012,0	8,37	7,78	12,96	34,22	25,11
20160102	29,22	24,83	1012,3	7,56	8,70	12,96	32,28	26,22
20160103	28,78	24,28	1012,7	7,56	8,33	14,82	33,22	25,39
20160104	29,44	25,44	1012,7	8,05	4,26	7,59	33,00	25,28
20160105	29,22	24,78	1012,1	8,05	8,89	14,82	32,61	26,78
20160106	30,17	24,39	1012,7	7,72	10,93	20,56	33,28	26,39
20160107	30,83	24,22	1012,4	7,72	6,11	11,11	34,61	26,28
20160108	29,94	24,50	1011,4	6,60	11,11	18,33	34,72	26,22
20160109	30,06	24,44	1010,9	8,21	4,82	11,11	34,22	25,89
20160110	29,22	24,50	1011,8	7,08	8,52	16,48	33,61	26,39
20160111	29,22	24,72	1011,8	8,37	7,78	20,56	32,22	26,11
20160112	29,89	24,50	1010,8	7,89	10,74	14,82	33,39	25,00
20160113	30,56	24,72	1010,4	8,69	12,96	18,33	33,00	26,61
20160114	30,94	25,11	1009,7	8,69	13,70	20,56	33,61	25,50
20160115	30,61	24,72	1009,0	9,17	7,96	18,33	34,22	25,61
20160116	30,33	23,89	1008,9	7,89	5,93	16,48	34,39	26,89
20160117	28,94	24,61	1008,7	6,60	7,59	18,33	34,00	25,78

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20160118	28,39	25,28	1009,2	5,63	10,93	20,56	33,22	25,72
20160119	28,56	24,94	1009,8	6,60	5,74	16,48	33,28	25,50
20160120	27,11	24,94	1011,1	5,63	10,19	20,56	31,78	25,39
20160121	27,61	24,56	1010,5	5,95	9,26	18,33	32,39	24,89
20160122	28,22	24,94	1010,9	5,63	7,04	18,33	31,89	24,61
20160123	26,61	24,39	1011,5	7,24	6,85	11,11	30,50	25,22
20160124	26,50	24,11	1012,4	7,89	5,56	14,82	29,61	25,00
20160125	26,28	23,94	1011,8	5,95	10,93	18,33	31,61	24,22
20160126	26,61	24,00	1011,4	6,60	7,96	22,22	30,72	24,22
20160127	26,67	23,44	1010,0	6,92	10,56	18,33	32,89	23,50
20160128	29,78	23,72	1009,6	8,05	12,59	18,33	34,39	23,89
20160129	29,83	24,50	1009,7	8,37	10,93	22,22	35,39	25,00
20160130	29,56	25,28	1008,4	7,24	14,45	27,78	34,39	26,00
20160131	28,28	24,83	1008,2	8,05	13,33	24,08	33,78	26,00
20160201	28,78	24,50	1008,1	7,89	16,30	33,52	33,39	25,89
20160202	27,67	24,61	1009,0	5,95	15,37	38,89	33,00	25,72
20160203	28,50	23,56	1009,1	8,69	8,52	18,33	31,39	25,22
20160205	26,61	24,50	1008,6	5,31	11,67	24,08	32,61	25,11
20160206	28,56	25,00	1008,5	8,05	9,45	14,82	32,78	24,72
20160207	27,06	24,56	1009,3	8,05	9,26	16,48	31,00	25,00
20160208	26,89	24,83	1008,8	7,40	4,82	11,11	31,61	24,89
20160209	27,78	24,83	1008,5	7,08	11,67	20,56	31,78	24,72
20160210	26,83	24,72	1009,1	4,83	8,33	11,11	31,39	25,11
20160211	26,72	24,39	1009,7	6,76	13,15	20,56	30,22	24,72
20160212	26,78	24,06	1010,5	6,60	7,96	14,82	29,22	24,61
20160213	28,89	24,78	1010,7	7,72	5,74	20,56	32,61	25,28
20160214	29,39	24,61	1010,2	7,89	9,07	14,82	33,22	25,78
20160215	27,39	25,06	1010,9	7,24	7,59	16,48	30,11	26,00
20160216	27,67	24,78	1012,2	6,76	5,74	14,82	30,61	25,39
20160217	28,11	24,06	1012,0	6,60	10,37	14,82	34,00	24,78
20160218	29,28	25,11	1010,6	6,92	4,07	16,48	33,22	24,61
20160219	29,72	25,00	1010,3	7,56	8,15	16,48	33,22	25,50
20160220	28,50	24,83	1011,2	7,08	8,33	16,48	32,28	26,28
20160221	27,00	24,78	1012,3	6,76	5,74	11,11	31,72	25,11
20160222	28,11	24,44	1012,3	5,63	9,26	27,78	33,11	24,50
20160223	26,78	24,17	1011,1	6,28	6,67	18,33	32,22	24,72
20160224	26,50	24,00	1010,7	6,60	8,70	14,82	32,61	24,11
20160225	27,67	24,67	1011,5	6,28	10,93	18,33	33,50	24,39

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20160226	26,78	24,17	1011,4	6,28	8,33	14,82	32,39	24,89
20160227	26,72	24,11	1011,2	7,56	11,48	16,48	32,28	24,50
20160228	28,17	24,39	1011,2	8,05	11,30	27,78	33,00	24,78
20160229	28,50	24,33	1011,5	8,53	16,85	27,78	32,61	25,00
20160302	27,78	24,56	1012,3	6,44	9,26	18,33	32,78	25,28
20160303	27,22	24,61	1011,3	6,12	6,48	9,45	33,00	25,00
20160304	28,33	24,94	1009,7	6,44	4,26	9,45	32,39	24,78
20160305	27,78	25,39	1010,5	4,83	7,41	16,48	31,50	20,39
20160306	27,44	24,61	1010,6	5,79	8,52	16,48	32,22	24,78
20160307	29,22	25,06	1009,9	6,60	4,82	12,96	33,00	24,72
20160308	29,33	25,22	1008,9	7,08	8,15	11,11	33,39	26,22
20160309	29,44	25,50	1009,8	6,12	6,30	14,82	32,89	25,89
20160310	29,83	25,89	1010,5	5,31	7,59	11,11	34,22	26,00
20160311	28,39	25,11	1010,6	7,56	10,56	18,33	33,00	24,72
20160312	29,00	24,72	1010,2	6,76	5,74	14,82	34,28	24,50
20160313	29,17	25,22	1010,5	6,76	8,33	16,48	33,61	26,00
20160314	28,61	24,94	1010,5	5,63	5,74	9,45	33,78	26,00
20160315	29,50	24,33	1009,0	7,72	7,59	14,82	33,61	25,78
20160316	29,06	25,22	1009,6	6,92	9,45	14,82	33,61	25,00
20160317	29,39	25,06	1009,7	7,56	8,33	12,96	33,00	26,22
20160318	28,22	25,22	1009,5	6,76	7,59	20,56	32,78	26,22
20160319	29,11	25,50	1009,6	5,95	6,30	20,56	32,78	25,89
20160320	30,17	25,61	1009,5	7,56	10,19	16,48	33,50	26,22
20160321	29,17	25,28	1008,8	6,92	7,96	16,48	33,78	26,39
20160322	29,11	25,06	1008,5	5,95	6,11	14,82	33,89	26,11
20160324	29,50	24,89	1011,8	5,95	5,56	14,82	33,61	25,50
20160325	28,94	25,28	1012,2	6,60	8,89	14,82	33,78	25,11
20160326	29,39	24,83	1013,0	7,40	8,15	11,11	33,22	24,78
20160327	29,39	24,72	1012,2	7,72	8,15	16,48	33,78	25,78
20160328	28,56	25,11	1011,8	6,92	5,74	11,11	32,28	25,78
20160329	28,44	25,11	1012,2	7,40	5,37	14,82	32,00	25,28
20160330	27,94	25,22	1012,5	6,76	6,85	14,82	32,61	25,61
20160331	28,00	24,67	1011,1	8,05	8,15	16,48	32,00	23,61
20160401	27,50	24,89	1010,8	7,08	5,74	16,48	30,22	24,89
20160403	29,50	25,28	1011,3	9,01	10,56	16,48	31,78	25,39
20160404	29,44	25,06	1010,4	8,05	15,93	27,78	33,00	24,78
20160405	29,11	25,44	1009,9	7,89	7,96	14,82	33,28	24,22
20160406	29,44	25,67	1010,7	8,05	7,41	22,22	32,22	24,61

Timestamp	TEMP	DEWP	SLP	VISIB	WDSP	MXSPD	MAX	MIN
20160407	29,11	25,83	1010,4	8,53	7,96	11,11	31,78	26,78
20160408	28,89	25,72	1010,1	8,05	7,04	11,11	33,39	25,50
20160409	27,67	25,11	1010,4	5,47	8,52	14,82	31,78	25,22
20160410	28,33	25,33	1010,3	6,60	7,78	11,11	31,39	24,89
20160411	29,39	25,72	1010,0	8,37	7,59	18,33	32,50	26,00
20160412	28,39	25,39	1009,5	7,89	7,59	16,48	31,61	25,61
20160413	28,11	25,00	1009,7	7,56	7,59	16,48	31,61	25,39
20160414	29,00	25,28	1010,2	7,24	8,33	16,48	33,00	25,61
20160415	27,44	25,39	1011,2	5,95	11,85	16,48	29,78	26,00
20160416	29,50	25,11	1010,6	9,01	8,52	14,82	32,39	26,00
20160417	28,56	25,33	1010,5	7,89	10,56	22,22	31,39	26,11
20160418	29,17	25,39	1010,0	7,72	7,41	18,33	32,78	26,22
20160419	29,06	24,94	1010,4	8,69	13,89	20,56	32,39	26,00
20160420	28,89	24,67	1010,6	8,05	10,74	20,56	34,39	25,89
20160421	29,11	24,56	1010,1	7,89	6,85	14,82	32,72	25,39
20160422	28,28	24,94	1010,4	8,05	10,37	16,48	31,11	25,50
20160423	29,89	24,78	1009,7	8,53	8,52	22,22	32,61	25,89
20160424	30,06	25,33	1009,4	8,69	15,19	20,56	32,89	27,28
20160425	30,50	24,78	1009,6	8,05	13,52	27,78	33,28	27,28
20160426	29,06	24,28	1010,5	8,05	8,15	20,56	32,50	25,78
20160427	29,72	24,11	1009,3	8,69	9,82	12,96	33,11	25,39
20160428	29,94	25,00	1009,7	7,89	8,70	16,48	33,39	26,78
20160429	29,83	25,06	1010,0	9,17	6,48	12,96	33,00	26,11
20160430	30,00	25,28	1009,8	7,56	11,85	25,93	32,61	26,61
20160501	29,33	25,44	1010,1	7,89	6,67	14,82	32,50	26,72
20160502	29,44	25,50	1010,3	8,69	10,74	20,56	32,50	26,61
20160503	28,83	25,78	1010,0	8,69	6,48	16,48	32,61	25,61
20160504	29,61	25,39	1009,0	8,05	7,96	14,82	32,61	26,11
20160505	30,06	25,50	1008,8	7,89	7,41	12,96	33,22	25,78
20160506	29,33	25,44	1008,3	6,28	5,37	11,11	32,61	26,89
20160507	28,78	25,50	1008,8	7,24	8,70	14,82	32,61	25,61
20160508	27,72	24,61	1009,5	6,76	7,41	12,96	32,89	23,61
20160509	28,89	25,89	1009,4	6,44	3,33	11,11	31,61	25,61
20160510	29,50	25,39	1008,8	7,72	6,67	16,48	32,39	25,78
20160511	28,89	25,61	1008,1	8,21	4,82	24,08	32,72	26,00
20160512	30,06	25,50	1007,8	8,05	9,63	24,08	33,22	26,50
20160513	30,72	25,61	1008,4	8,69	14,08	27,78	33,72	26,89
20160514	30,50	25,78	1009,9	8,05	12,96	18,33	33,39	26,78

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20160515	30,00	25,50	1008,5	8,21	12,41	20,56	33,22	26,00
20160516	29,28	25,22	1008,5	6,92	7,04	14,82	33,39	25,61
20160517	29,61	24,83	1009,1	7,56	9,45	12,96	33,22	25,39
20160518	30,11	24,28	1009,0	7,56	8,89	14,82	33,22	26,00
20160519	30,06	25,50	1008,6	7,89	9,07	16,48	33,22	26,50
20160520	28,89	25,39	1008,4	7,72	10,19	16,48	32,22	25,61
20160521	28,89	25,11	1010,3	8,53	6,48	20,56	32,00	25,22
20160522	29,89	25,06	1010,8	8,53	12,41	18,33	32,61	26,50
20160523	29,67	24,72	1011,5	8,05	7,78	14,82	33,00	26,28
20160525	30,11	24,56	1010,0	7,72	25,37	79,45	32,61	27,39
20160526	30,39	24,83	1010,9	9,01	14,26	22,22	32,72	26,89
20160527	28,11	24,22	1011,9	8,37	12,59	22,22	32,22	26,11
20160528	28,94	25,28	1010,6	9,01	7,96	18,33	32,00	26,00
20160529	29,06	25,50	1010,0	7,40	10,93	18,33	31,78	26,39
20160530	26,61	24,44	1011,6	5,95	9,07	24,08	31,22	24,00
20160531	27,61	24,00	1011,9	6,44	4,63	9,45	32,11	23,89
20160601	28,56	25,06	1010,9	7,72	7,96	16,48	31,78	25,78
20160602	29,56	24,22	1010,7	8,05	10,56	20,56	32,00	25,61
20160603	29,22	23,83	1011,3	8,69	8,33	14,82	31,89	26,11
20160604	29,39	23,94	1011,3	7,56	8,15	12,96	33,00	25,28
20160605	29,28	24,67	1011,7	6,92	4,82	14,82	32,28	25,89
20160606	29,11	24,83	1011,9	8,05	9,63	14,82	32,50	25,89
20160607	29,33	25,44	1011,9	8,05	5,56	16,48	32,61	25,39
20160608	29,56	24,28	1011,2	7,72	10,74	18,33	32,39	25,78
20160609	29,17	24,72	1012,0	7,40	10,37	24,08	32,28	26,22
20160610	29,61	24,50	1012,1	8,85	13,70	18,33	32,11	25,72
20160611	29,78	24,11	1012,2	8,69	12,59	22,22	32,50	26,50
20160612	29,72	24,39	1011,5	8,05	11,11	16,48	32,39	26,39
20160613	29,00	24,78	1010,7	6,92	8,33	18,33	32,39	26,22
20160614	28,11	24,61	1010,4	7,40	6,67	16,48	31,00	25,50
20160616	29,44	24,39	1009,0	8,05	11,85	20,56	32,39	25,28
20160617	28,39	25,22	1008,6	6,92	6,67	16,48	32,22	25,50
20160618	27,39	24,50	1009,5	7,40	10,93	20,56	32,22	24,39
20160619	26,22	23,61	1010,0	6,76	6,30	22,22	31,39	23,78
20160620	28,33	23,83	1009,2	7,56	4,44	9,45	31,39	23,28
20160621	28,78	24,83	1009,8	7,56	8,33	11,11	31,22	24,50
20160622	27,89	23,83	1010,5	7,40	6,67	12,96	31,39	24,39
20160623	28,17	24,50	1010,3	6,92	7,59	11,11	31,72	24,00

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20160624	28,83	24,11	1010,0	6,92	7,41	14,82	32,39	25,00
20160625	28,94	23,89	1010,9	7,24	9,63	16,48	31,78	25,61
20160626	28,72	23,72	1010,6	7,56	9,45	18,33	32,22	25,11
20160627	28,22	24,33	1009,8	7,40	6,30	12,96	32,39	25,22
20160628	27,67	24,00	1010,7	6,76	12,78	20,56	30,89	25,00
20160629	27,50	25,28	1011,0	6,60	6,85	25,93	30,00	24,61
20160630	26,11	24,39	1011,5	8,05	5,74	12,96	28,61	24,61
20160701	27,78	24,22	1011,5	7,72	8,52	14,82	31,50	24,50
20160702	28,89	24,89	1012,1	7,89	10,00	18,33	31,61	25,00
20160703	29,06	24,72	1011,6	8,21	14,63	24,08	31,78	25,22
20160704	29,56	24,00	1010,9	9,17	16,85	20,56	31,61	25,50
20160705	29,06	24,11	1010,8	9,50	13,33	18,33	31,50	26,22
20160706	29,11	24,28	1010,2	8,69	15,00	25,93	31,50	25,89
20160707	28,67	23,67	1010,4	8,69	11,30	20,56	31,61	25,00
20160708	29,28	22,94	1010,5	8,21	13,89	18,33	32,22	24,22
20160709	29,22	24,17	1010,8	8,53	12,04	20,56	32,22	26,00
20160712	28,61	24,94	1010,0	7,89	6,85	16,48	31,72	25,89
20160713	28,61	24,89	1009,9	8,05	9,82	20,56	31,50	25,39
20160714	28,61	25,00	1009,6	7,56	12,59	20,56	31,61	25,11
20160715	28,28	24,78	1009,7	6,44	12,04	27,78	31,11	25,89
20160716	27,67	24,67	1010,9	6,60	11,67	18,33	29,61	25,61
20160717	27,89	25,00	1010,9	7,72	9,45	20,56	31,39	25,39
20160718	26,89	24,28	1009,9	6,92	7,78	20,56	30,39	24,39
20160719	28,56	23,50	1009,1	8,05	6,85	12,96	31,89	24,39
20160720	26,28	24,39	1010,2	5,95	6,11	18,33	30,50	24,61
20160723	28,28	24,56	1009,5	6,44	8,33	18,33	31,39	25,61
20160724	28,00	23,17	1010,0	7,40	8,33	18,33	31,39	25,00
20160725	28,11	22,78	1010,3	6,92	10,37	18,33	31,39	24,11
20160726	28,28	22,72	1010,4	7,24	8,52	16,48	31,50	24,11
20160727	28,72	23,00	1010,5	7,89	10,74	16,48	31,72	24,72
20160728	29,00	23,33	1010,7	6,92	9,82	18,33	31,22	24,72
20160729	28,78	22,61	1010,0	6,44	7,04	18,33	31,61	25,89
20160731	27,67	23,11	1010,3	7,40	20,74	70,56	31,00	24,00
20160801	27,78	22,39	1009,8	7,08	8,52	16,48	31,39	24,78
20160802	27,39	20,89	1009,3	7,72	9,45	14,82	31,39	23,78
20160803	28,00	22,39	1009,6	7,24	7,78	14,82	31,00	23,61
20160804	28,72	23,72	1010,4	7,72	11,11	20,56	31,61	24,50
20160805	28,39	24,28	1010,8	7,56	13,70	22,22	31,39	24,39

Timestamp	TEMP	DEWP	SLP	VISIB	VDSP	MXSPD	MAX	MIN
20160808	29,00	24,89	1011,3	9,82	9,63	18,33	31,39	25,50
20160809	28,72	23,61	1011,2	8,85	13,70	29,45	31,39	25,72
20160810	28,00	23,00	1011,2	5,95	11,11	27,78	32,00	24,00
20160811	28,39	22,50	1011,7	7,56	8,89	18,33	32,22	23,50
20160812	29,00	23,72	1010,9	7,89	11,85	18,33	32,50	24,39
20160813	28,00	23,94	1010,6	7,56	15,93	25,93	31,11	25,50
20160814	28,67	22,94	1010,4	8,21	11,48	18,33	31,39	25,00
20160815	28,56	23,17	1010,7	8,53	15,56	24,08	31,39	24,89
20160816	28,94	23,67	1010,6	8,69	10,37	18,33	31,78	25,00
20160817	28,67	23,72	1010,0	8,05	13,33	20,56	31,78	25,28
20160818	28,22	22,72	1009,0	7,56	11,30	22,22	32,22	24,50
20160819	27,89	20,83	1009,4	8,05	13,70	27,78	31,61	24,22
20160820	27,72	20,78	1009,6	7,56	12,59	18,33	31,89	23,39
20160821	27,56	20,33	1009,5	7,24	8,52	16,48	31,72	23,00
20160822	27,33	20,33	1009,1	7,56	9,26	18,33	31,39	23,11
20160823	27,28	21,00	1008,3	7,56	12,22	21,67	31,22	22,50
20160824	27,50	21,61	1008,4	7,40	12,04	27,78	31,00	23,28
20160825	28,11	21,89	1009,9	7,40	10,74	18,33	31,22	23,61
20160826	28,67	23,39	1011,0	7,56	7,96	14,82	30,78	24,28
20160827	29,17	24,83	1011,5	8,05	18,89	27,78	31,78	25,11
20160828	29,22	24,17	1011,0	8,53	12,96	22,22	32,00	25,72
20160829	29,44	24,28	1010,8	8,69	17,22	20,56	31,89	26,78
20160830	29,50	24,11	1011,1	8,53	18,71	27,78	32,22	22,39
20160831	29,56	23,39	1012,1	8,69	12,59	20,56	32,61	26,22
20160901	29,50	23,50	1010,5	6,92	11,30	20,56	33,28	25,28
20160902	29,39	22,89	1010,1	7,08	11,48	16,48	33,39	25,22
20160903	28,67	23,50	1010,2	6,76	10,37	20,56	33,00	25,00
20160904	29,11	23,50	1011,5	7,56	11,48	18,33	32,39	24,72
20160905	29,44	23,78	1011,9	8,37	19,08	25,93	31,89	26,39
20160906	29,67	24,00	1012,2	8,85	12,78	22,22	32,50	26,39
20160907	29,39	23,17	1011,7	7,72	15,00	25,93	32,61	25,78
20160908	29,56	22,39	1011,8	8,05	14,26	25,93	33,00	25,00
20160909	29,61	22,67	1011,8	7,56	15,74	24,08	32,78	25,50
20160910	29,39	22,67	1010,5	6,76	12,78	24,08	33,39	24,89
20160911	29,11	21,50	1009,9	7,89	12,04	25,93	33,00	25,00
20160912	28,50	22,17	1010,7	7,89	11,11	25,93	33,00	25,22
20160913	28,72	22,56	1010,7	7,56	8,33	14,82	32,61	24,22
20160914	29,17	22,50	1010,4	8,05	11,48	18,33	33,00	24,39

Timestamp	TEMP	DEWP	SLP	VISIB	WDSP	MXSPD	MAX	MIN
20160916	29,44	24,33	1009,2	8,53	14,26	22,22	32,39	25,39
20160917	29,11	24,17	1009,8	8,37	14,08	27,78	32,72	26,61
20160918	29,67	24,67	1009,5	9,01	11,11	25,93	32,78	27,39
20160919	29,50	24,56	1008,8	8,21	6,30	14,82	32,61	26,50
20160922	29,50	24,28	1010,1	8,69	9,07	16,48	33,22	26,00
20160923	28,44	24,28	1008,8	6,92	11,11	18,33	33,00	23,22
20160924	26,78	24,22	1008,7	6,92	7,96	18,33	31,89	22,72
20160927	28,06	24,33	1009,3	8,05	9,63	14,82	32,50	20,72

LAMPIRAN B

DATA VARIABEL TURUNAN, CURAH HUJAN DAN KELAS INTENSITAS CURAH HUJAN

**Tabel B-1 Data Variabel Turunan, Curah Hujan Dan Kelas
Intensitas Curah Hujan**

TD	RT	RH	B L N	PRC P	Class	TD	RT	RH	B L N	PRC P	Class
4,67	9,89	75,88	1	0	tidak	5	8,72	74,31	2	0	tidak
4,28	7,61	77,68	1	0	tidak	4,89	7,5	74,77	2	0	tidak
4,45	8,22	76,87	1	0	tidak	4	6,78	78,88	2	0	tidak
3,23	10,3	82,49	1	82,04	lebat	4,16	8	78,12	2	0,76	ringan
1,72	6,61	90,19	1	0	tidak	4,33	8	77,37	2	0	tidak
3,17	9,61	82,84	1	76,96	lebat	2,84	8,61	84,38	2	38,1	sedang
2,5	9,39	86,17	1	27,94	sedang	2,78	7,78	84,71	2	20,07	sedang
2,27	6,72	87,31	1	0	tidak	3,89	8,22	79,36	2	13,72	ringan
2,39	7,28	86,63	1	2,03	ringan	3,89	6,61	79,39	2	0	tidak
5,83	10,4	70,74	1	0	tidak	4,77	5,78	75,33	2	0	tidak
3,77	9,5	79,93	1	0	tidak	4,94	5,72	74,63	3	2,03	ringan
3	9,61	83,61	1	53,09	lebat	3	5	83,6	3	0,76	ringan
3,78	7	79,84	1	3,05	ringan	4,06	17,6	78,53	3	0	tidak
4,89	8,5	74,77	1	0	tidak	4,61	7,61	76,07	3	0	tidak
5,5	8,17	72,04	1	0	tidak	4,11	7,22	78,36	3	0	tidak
3,61	6,89	80,62	1	26,67	sedang	5,17	10,8	73,52	3	0	tidak
2,95	5,28	83,81	1	4,06	ringan	6,22	8,72	69,04	3	0	tidak
4,17	7,06	77,99	1	6,1	ringan	6,06	8	69,89	3	0	tidak
2,88	8,28	84,17	1	0,76	ringan	5,33	8,94	72,92	3	0	tidak
4,84	8,39	74,99	2	0	tidak	4,67	9,5	75,81	3	7,11	ringan
3,94	6,39	79,07	2	0,76	ringan	4	6,78	78,88	3	0	tidak
4,56	5,22	76,21	2	0	tidak	4,39	7,5	77,07	3	0	tidak
4,89	6,89	74,82	2	13,97	ringan	5,05	8,39	74,18	3	0	tidak
2,33	5,72	86,95	2	4,06	ringan	3,83	6,78	79,57	3	0,76	ringan
2,22	6,78	87,58	2	53,09	lebat	2,83	7,33	84,42	3	0,76	ringan
2,39	7,06	86,65	2	42,93	sedang	2,89	8,11	84,24	3	0	tidak
3,95	6,06	78,99	2	0	tidak	4,05	8	78,68	3	0	tidak
5,5	7,28	72,18	2	0	tidak	4,78	6,22	75,41	3	0	tidak
4,77	9,11	75,4	2	2,03	ringan	3,34	5,61	82,08	4	0	tidak
3,17	9	82,84	2	0	tidak	3,17	7,39	82,83	4	10,92	ringan
2,89	8,06	84,24	2	0	tidak	5,17	7,33	73,71	4	0	tidak
3,78	10,2	79,93	2	4,83	ringan	3,22	6,61	82,6	4	0,51	ringan
4,56	8,11	76,39	2	0	tidak	3,95	6,33	79,21	4	0	tidak

B-2

TD	RT	RH	B L N	PRC P	Class
4,5	8,61	76,6	4	44,96	sedang
3,44	8	81,55	4	0	tidak
2,83	7,39	84,53	4	16	ringan
3,55	7,83	81,06	4	0	tidak
3,62	5,89	80,68	4	0	tidak
2,72	6,61	85,04	4	0	tidak
3,44	7,11	81,55	4	0	tidak
5,11	7,89	73,93	4	0,76	ringan
4,66	7,39	75,9	4	0	tidak
3,67	8,78	80,42	4	7,11	ringan
4,88	7,72	74,91	4	0	tidak
3	7,22	83,69	4	2,03	ringan
4,33	7,11	77,5	4	0	tidak
2,05	3,39	88,49	4	2,03	ringan
3,28	7,11	82,33	4	2,03	ringan
2,77	6,44	84,78	4	0,76	ringan
4,73	7,39	75,61	4	0	tidak
4,28	7,78	77,66	4	0	tidak
6,34	7,22	68,71	4	0	tidak
3,55	6,5	80,97	5	0,76	ringan
4,16	6,5	78,2	5	0	tidak
5,22	8	73,42	5	0	tidak
6	7,61	69,98	5	0	tidak
5,56	7,72	71,96	5	0	tidak
5,61	8	71,69	5	0	tidak
5,78	9,22	71,03	5	0	tidak
6,16	8,39	69,49	5	0	tidak
6	8,22	70,2	5	0	tidak
5,66	6,61	71,64	5	0	tidak
3,78	6,22	80,11	5	6,86	ringan
3,94	5,39	79,29	5	0	tidak
5,44	7,28	72,56	5	0	tidak
2,72	7,39	85,11	5	0,76	ringan
3,78	6,39	80,04	5	0	tidak
3,83	7	79,79	5	0	tidak
5,34	7	72,99	5	1,78	ringan
5,17	7,11	73,69	5	0,76	ringan
5	7,61	74,41	5	0	tidak
5,5	7,22	72,36	5	0	tidak
6,84	6,61	66,78	5	0	tidak
6,61	9	67,39	5	0	tidak
5,45	8,61	72,18	6	0	tidak
6,23	7,78	69,15	6	0,76	ringan
3,5	7	81,2	6	32	sedang
3	6,39	83,73	6	1,02	ringan
4,28	6,44	77,63	6	0	tidak
5,5	9,39	72,1	6	0	tidak
5,89	8,39	70,34	6	0	tidak
7,94	9,39	62,28	6	0	tidak
5,33	7,39	72,8	6	0	tidak
5,55	7,39	71,8	6	0	tidak
4,78	8,39	75,16	6	0	tidak
5,22	8,72	73,32	6	0	tidak
6,78	7,89	66,71	6	0	tidak
6,56	8,83	67,51	6	0	tidak
6,34	8,78	68,44	6	0	tidak
6,45	6,17	68,25	6	0	tidak
3,89	3,39	79,22	6	2,03	ringan
5,5	7,28	72,12	6	0	tidak
6,11	7,39	69,51	6	0	tidak
4,89	7,78	74,75	6	0	tidak
6,44	8,61	68,17	6	0	tidak
7,84	8,61	62,49	6	0	tidak
7,61	10,3	63,27	6	0	tidak
6,22	11,7	68,67	6	0	tidak
7,06	10	65,38	6	0	tidak
6,61	9,61	67,25	6	0	tidak
6,66	8,44	67,12	6	0	tidak
6	9	69,71	6	0	tidak
5,56	8	71,74	7	0	tidak
5,39	7,28	72,45	7	0	tidak
7,11	7,44	65,4	7	0	tidak
6,39	8,89	68,35	7	0	tidak
6,44	8,61	68,13	7	0	tidak
7,11	8	65,4	7	0	tidak
8	7	62,21	7	0	tidak
6,06	8	69,68	7	0	tidak
7,06	8,5	65,51	7	0	tidak
7,73	9,28	62,7	7	0	tidak
8,11	10	61,35	7	0	tidak

TD	RT	RH	B L N	PRC	Class
6,72	9,78	67	7	0	tidak
7,17	9,89	65,25	7	0	tidak
6,22	9,89	68,78	7	0	tidak
8,5	9,39	59,92	7	0	tidak
6,83	6,5	66,55	7	0	tidak
7,17	6,11	65,23	7	0	tidak
7,11	7	65,51	7	0	tidak
6,05	8	69,68	7	0	tidak
6,56	8	67,37	7	0	tidak
9,28	8,5	56,92	7	0	tidak
8,67	9,61	59,22	7	0	tidak
6,45	10,1	67,89	7	0	tidak
8,22	10,6	61,31	7	0	tidak
5,78	8,5	70,82	7	0	tidak
6,22	7,17	69,03	7	0	tidak
5,44	7,78	72,34	7	0	tidak
6,5	8	67,84	7	0	tidak
7,28	9,11	64,84	7	0	tidak
5,78	9,61	70,73	8	0	tidak
6,44	9,5	67,86	8	0	tidak
6,56	9,39	67,57	8	0	tidak
7,28	8,78	64,58	8	0	tidak
7,11	9,61	65,17	8	0	tidak
6,11	9,5	69,34	8	0	tidak
5,89	10,7	70,4	8	0	tidak
7,28	8,83	64,56	8	0	tidak
6,39	9,28	68,21	8	0	tidak
7,05	10	65,53	8	0	tidak
8,77	9,83	59,05	8	0	tidak
9,55	4,22	56,38	8	0	tidak
5,44	8,5	72,19	8	0	tidak
9,05	9	57,98	8	0	tidak
7,11	9,11	65,19	8	0	tidak
7,28	10,2	64,57	8	0	tidak
7,06	8,61	65,3	8	0	tidak
8,05	8,78	61,72	8	0	tidak
7,5	8,67	63,6	8	0	tidak
7,94	9	61,94	8	0	tidak
8,27	10	60,68	8	0	tidak
10,19	9,83	54,06	8	0	tidak

TD	RT	RH	B L N	PRC	Class
5,94	9,67	70,01	8	0	tidak
7,44	9,22	64,3	8	0	tidak
6,05	7,83	69,82	8	0	tidak
8,33	7	60,92	8	0	tidak
5,72	7,5	71,18	8	0	tidak
8,38	9,22	60,72	8	0	tidak
6	8	69,85	8	0	tidak
10,78	8,61	52,78	8	0	tidak
6,83	6,22	66,64	8	0	tidak
7,78	7,06	62,94	9	0	tidak
6,66	7,78	67,44	9	0	tidak
7,77	7	62,9	9	0	tidak
6,45	9,39	68,01	9	0	tidak
8,16	8	61,66	9	0	tidak
7,27	10	64,76	9	0	tidak
8,5	8	60,32	9	0	tidak
9,88	10,7	55,47	9	0	tidak
7,45	10,3	63,93	9	0	tidak
9,55	10,9	56,38	9	0	tidak
8,55	10,6	59,93	9	0	tidak
7,05	9,5	65,74	9	0	tidak
6,56	11	67,7	9	0	tidak
8,22	9,72	61,21	9	0	tidak
9,11	11,2	57,94	9	0	tidak
5,34	11	72,66	9	0	tidak
8,5	9,39	60,36	9	0	tidak
7,61	7,28	63,65	9	0	tidak
7,22	6,39	65,26	9	0	tidak
8	8	62,21	9	0	tidak
6,06	8,39	69,79	9	0	tidak
8,11	9	61,62	9	0	tidak
7,95	9,94	62,37	9	0	tidak
7,72	8,61	63,24	9	0	tidak
8,89	10,4	58,98	9	0	tidak
7,28	9,11	64,9	9	0	tidak
8,22	9	61,32	9	0	tidak
7,89	10	62,52	9	0	tidak
8,22	9	61,39	9	0	tidak
7,39	9,22	64,39	9	0	tidak
8,72	9,78	59,43	10	0	tidak

B-4

TD	RT	RH	B L N	PRC P	Class	TD	RT	RH	B L N	PRC P	Class
7,84	9,61	62,9	10	0	tidak	4,28	6,28	77,62	11	0,51	ringan
7,78	9	63,05	10	0	tidak	4,67	8,11	75,98	11	0	tidak
7,56	9,11	63,9	10	0	tidak	4,05	8,78	78,63	11	0,76	ringan
7,5	10,3	64,11	10	0	tidak	2,44	4,5	86,45	11	0,51	ringan
8,11	9,61	61,84	10	0	tidak	4,5	7,11	76,55	11	2,03	ringan
9	10	58,71	10	0	tidak	4,22	8,39	77,9	11	1,27	ringan
7,72	9,89	63,23	10	0	tidak	5	6,5	74,27	11	0	tidak
9,34	10,6	57,41	10	0	tidak	7,39	11	64,66	12	0	tidak
7,45	9,72	64,31	10	0	tidak	7,94	9,61	62,25	12	0	tidak
9,72	8,61	56,25	10	0	tidak	7,33	10,2	64,45	12	0	tidak
8,05	8,61	62,14	10	0	tidak	9,89	11,2	55,45	12	0	tidak
6,45	8,89	68,33	10	0	tidak	7,67	12,2	63,17	12	0	tidak
5,89	8,61	70,67	10	0	tidak	10,5	10,8	53,63	12	0	tidak
5,56	8,17	72,01	10	0	tidak	9,72	10,6	56,23	12	0	tidak
5,22	7,61	73,4	10	0	tidak	7,78	11,1	63,13	12	0	tidak
5,56	7,33	71,96	10	0	tidak	5	8,83	74,37	12	0	tidak
5	8,22	74,38	10	0	tidak	3,11	7,67	83,15	12	18,03	ringan
6,22	9	69,2	10	0	tidak	4,11	9,22	78,42	12	51,05	lebat
6,61	7	67,55	10	0	tidak	2,78	4,72	84,76	12	0	tidak
6,83	9	66,56	10	0	tidak	3,05	6,11	83,44	12	0	tidak
5,33	5	72,81	10	0	tidak	3,23	7,39	82,53	12	41,91	sedang
4,39	6	77,1	10	0	tidak	3,77	6,5	79,99	12	2,03	ringan
2,67	8,22	85,33	10	5,08	ringan	4,66	6,61	75,96	12	0	tidak
4,95	7	74,65	10	0,76	ringan	3,33	6	82,12	12	0	tidak
4,22	6,22	77,94	11	6,1	ringan	4,34	6	77,29	12	0	tidak
3,67	8	80,46	11	1,02	ringan	5,67	8,78	71,46	12	0	tidak
5,67	8,44	71,57	11	0	tidak	8,22	10,2	61,77	12	0	tidak
4,61	8,56	76,13	11	17,02	ringan	7,89	11	62,71	12	0	tidak
5,78	8,61	71,12	11	0	tidak	6,67	7,78	67,47	12	0	tidak
5	7,39	74,56	11	0	tidak	5,83	7,22	70,92	12	0	tidak
5	8,83	74,54	11	0	tidak	5,55	9,22	72,08	12	0	tidak
3,17	6,78	82,87	11	1,02	ringan	6,45	8,61	68,45	12	0	tidak
5,39	8,33	72,8	11	0	tidak	6,55	9	67,95	12	0	tidak
4,45	8,11	76,91	11	37,08	sedang	6,72	8,5	67,29	1	0	tidak
3,45	7	81,57	11	0,76	ringan	5	7,89	74,36	1	0,76	ringan
5	9	74,48	11	0	tidak	4,89	7,5	74,91	1	0	tidak
3,94	6,72	79,29	11	0	tidak	2,61	7	85,59	1	0	tidak
3,61	7,78	80,76	11	0	tidak	2,66	8	85,33	1	8,89	ringan
2,83	9,39	84,49	11	7,87	ringan	1,72	8,61	90,19	1	11,94	ringan
4,77	7,78	75,33	11	0	tidak	5,06	8,22	74,01	1	0	tidak

TD	RT	RH	B L N	PRC P	Class
3,67	4,22	80,24	1	0	tidak
5,39	6	72,6	1	0	tidak
6,06	6,61	69,81	1	0	tidak
5,61	9,5	71,74	1	0	tidak
2,89	6,78	84,23	1	8,89	ringan
2,72	8,5	85,04	1	45,97	sedang
3,78	8,11	79,89	1	0,76	ringan
5,45	7,11	72,51	1	0	tidak
3,12	6,44	83,2	1	0	tidak
5,22	7	73,62	1	0	tidak
3,39	10,1	81,74	1	23,11	sedang
3,11	8,39	83,18	1	25,91	sedang
3,45	7,83	81,51	1	2,03	ringan
3,72	8,5	80,23	1	2,03	ringan
3,78	6,5	79,72	2	0,76	ringan
5,05	7,94	74,07	2	5,08	ringan
2,89	7,78	84,18	2	14,99	ringan
4,39	9	77,02	2	0	tidak
3,34	7	81,88	2	8,89	ringan
3,22	7,83	82,55	2	3,05	ringan
5,34	8,61	72,8	2	0	tidak
4,95	7,72	74,47	2	0	tidak
6	8,39	70,1	2	0	tidak
5,61	8,33	71,63	2	0,76	ringan
7,28	9,89	64,94	2	0	tidak
4,11	9,89	78,29	2	32,77	sedang
4,72	10	75,46	2	3,81	ringan
4,55	8,39	76,26	2	2,03	ringan
6,05	8,11	69,75	2	0	tidak
5,45	5,5	72,42	2	0	tidak
2,84	7	84,44	2	2,03	ringan
5,16	7,89	73,66	2	0	tidak
3,27	8,11	82,38	2	4,06	ringan
3,84	8,11	79,65	3	0	tidak
3,27	8	82,37	3	27,94	sedang
5,17	9,78	73,64	3	0	tidak
1,94	9,61	89,03	3	71,12	lebat
4	9,61	78,85	3	68,07	lebat
3,61	8,28	80,67	3	0	tidak
3,84	7	79,64	3	9,91	ringan

TD	RT	RH	B L N	PRC P	Class
3,61	6,83	80,7	3	0	tidak
2,27	9,39	87,32	3	53,09	lebat
3,23	7,83	82,48	3	18,03	ringan
4,27	9,06	77,64	3	0	tidak
5,61	8,89	71,74	3	0,51	ringan
4,16	6,72	78,12	3	0	tidak
2,94	7,89	84,03	3	7,11	ringan
3,89	8,5	79,34	3	5,84	ringan
3,56	10	80,88	3	21,08	sedang
5,67	9,5	71,48	3	0	tidak
5,39	9,89	72,62	3	7,87	ringan
5,89	7,89	70,56	3	0	tidak
2,5	8,89	86,21	3	4,06	ringan
2,39	8	86,71	3	1,27	ringan
3,44	7,61	81,55	3	0	tidak
4,16	7,39	78,14	3	11,68	ringan
3,28	8,56	82,35	3	0	tidak
3,39	7,89	81,81	3	39,88	sedang
2,27	7,22	87,4	3	7,87	ringan
2,61	7	85,63	3	23,88	sedang
3,5	7,5	81,3	4	0	tidak
3,83	7,56	79,77	4	0	tidak
3,11	7,61	83,16	4	23,11	sedang
3,95	7,28	79,17	4	0	tidak
4,06	8,61	78,73	4	0	tidak
3,67	7,83	80,56	4	3,05	ringan
3,39	7,5	81,85	4	0	tidak
3,34	8,61	82,02	4	3,05	ringan
3,16	7,39	82,9	4	27,69	sedang
3,5	5	81,28	4	0	tidak
2,94	7,5	83,98	4	23,88	sedang
3,33	7	82,03	4	0,76	ringan
4,45	8,5	76,69	4	0	tidak
7,12	8,33	65,51	4	0	tidak
6,11	7,61	69,5	4	0	tidak
6,72	7,5	67,11	4	0	tidak
6,95	8,39	66,23	4	0	tidak
3,78	7	80,07	4	0	tidak
5,22	7,72	73,53	4	0	tidak
5	9,39	74,36	4	0	tidak

B-6

TD	RT	RH	B L N	PRC P	Class	TD	RT	RH	B L N	PRC P	Class
5,83	8,83	70,78	4	0	tidak	6,28	8	68,89	6	0	tidak
8,78	8,33	59,35	4	0	tidak	5,72	9	71	6	0	tidak
3,89	7,83	79,51	5	0	tidak	8,11	9,78	61,42	6	0	tidak
3,62	8	80,75	5	3,05	ringan	7,78	8,89	62,6	6	0	tidak
5,45	7,39	72,54	5	7,87	ringan	6,61	9,78	67,3	6	0	tidak
5,95	8,11	70,46	5	0	tidak	4	5,78	78,89	6	34,04	sedang
6,17	7,61	69,39	5	0	tidak	5,89	6	70,6	6	0	tidak
7,05	7,78	65,9	5	0	tidak	5,33	5,61	72,82	6	0	tidak
6,56	8,39	67,6	5	0	tidak	5,55	9,67	71,68	6	0	tidak
7	9,61	65,74	5	0	tidak	6,33	8,44	68,53	6	0	tidak
6,5	9,5	67,78	5	0	tidak	5,62	8,78	71,48	6	0	tidak
7,61	8,61	63,59	5	0	tidak	5,83	8,22	70,62	6	0	tidak
6,66	7,61	67,32	5	0	tidak	5,67	8,78	71,32	6	0	tidak
6,16	12,2	69,38	5	0	tidak	9,56	11,2	55,91	7	0	tidak
5,83	7,56	70,81	5	0	tidak	7	10,7	65,38	7	0	tidak
5,95	7,5	70,42	5	0	tidak	8,05	11	61,04	7	0	tidak
4,44	8	76,97	5	0	tidak	8,83	10	58,24	7	0	tidak
3,67	8	80,47	5	26,92	sedang	10,2	10	53,64	7	0	tidak
4,28	5,72	77,64	5	0	tidak	7,78	9	62,6	7	0	tidak
5,94	6,78	70,4	5	0	tidak	5,56	7	71,71	7	0	tidak
7,22	8	65,27	5	0	tidak	5,11	6	73,68	7	0	tidak
7,72	7	63,35	5	0	tidak	5,78	6	70,91	7	0	tidak
5,83	9,22	70,65	5	0	tidak	5,33	6	72,94	7	0	tidak
7,38	9	64,45	5	0	tidak	4,89	6,89	74,81	7	0	tidak
5,84	8	70,8	5	0	tidak	5,44	5,5	72,42	7	0	tidak
5,27	8	73,34	5	0	tidak	4,39	7,89	77,02	7	0	tidak
5,5	6	72,26	5	0	tidak	5	7,11	74,3	7	0	tidak
5,72	4	71,41	5	0	tidak	5,22	7,28	73,23	7	0	tidak
5,61	7	71,86	5	0	tidak	4,11	6	78,25	7	3,05	ringan
5	6	74,52	6	0	tidak	4,61	7	76,04	7	0	tidak
6,05	6	70,01	6	0	tidak	4,78	5,17	75,31	7	0	tidak
4,45	7	76,85	6	0	tidak	4,94	5,89	74,51	7	0	tidak
4,61	7	76,06	6	0	tidak	6,17	7,61	69,28	7	0	tidak
4,95	6	74,62	6	0	tidak	6,83	7,61	66,51	7	0	tidak
5,61	7	71,81	6	0	tidak	6,94	8,06	66,06	7	0	tidak
5,17	7	73,59	6	0	tidak	7,39	9	64,24	7	0	tidak
3,17	8,22	82,86	6	35,05	sedang	5,39	9,5	72,53	7	0	tidak
3,45	6	81,45	6	1,02	ringan	5,67	9	71,23	7	0	tidak
4,33	6,5	77,43	6	2,03	ringan	6,83	9,39	66,34	7	0	tidak
5,45	7,06	72,53	6	0	tidak	6,94	9	65,8	7	0	tidak

TD	RT	RH	B L N	PRC	Class
7,27	10,1	64,46	8	0	tidak
6,89	11,5	65,97	8	0	tidak
6,11	10	68,91	8	0	tidak
6,78	10,1	66,18	8	0	tidak
7,84	10,3	62,05	8	0	tidak
9,17	10,6	57,39	8	0	tidak
6,11	8,72	69,23	8	0	tidak
5,83	10	70,53	8	0	tidak
5,77	9,39	70,81	8	0	tidak
6,11	9,11	69,5	8	0	tidak
6	10	69,76	8	0	tidak
6,67	9,39	66,97	8	0	tidak
7,44	8,33	63,87	8	0	tidak
7,39	9,72	64,15	8	0	tidak
5,61	8,39	71,5	8	0	tidak
6,56	10	67,48	8	0	tidak
8,5	9,39	59,7	8	0	tidak
8,5	10,8	59,7	8	0	tidak
6,84	8,78	66,27	8	0	tidak
7,5	8,39	63,82	8	0	tidak
6,5	9,22	67,54	8	0	tidak
7,44	9,56	64	8	0	tidak
7,22	10,1	64,86	8	0	tidak
6,22	9,39	68,67	8	0	tidak
7,62	9,39	63,21	8	0	tidak
6,28	9,61	68,6	8	0	tidak
5,89	10,2	70,04	8	0	tidak
8,56	10,1	59,62	8	0	tidak
6,11	9,78	69,41	8	0	tidak
5,05	7,89	74,1	8	0	tidak
5,72	8,28	71,26	8	0	tidak
7,05	8,83	65,66	9	0	tidak
5,61	7	71,56	9	0	tidak
5,95	6,89	70,17	9	0	tidak
6,17	6,83	69,31	9	0	tidak
5,55	8	71,8	9	0	tidak
8,73	10,4	59,25	9	0	tidak
7,17	9,78	65,25	9	0	tidak
6,61	10	67,32	9	0	tidak
6,45	9	68,08	9	0	tidak

TD	RT	RH	B L N	PRC	Class
7,22	9,39	64,85	9	0	tidak
6,56	11	67,48	9	0	tidak
7,22	10,6	65,02	9	0	tidak
7,5	9,78	63,9	9	0	tidak
6,56	10	67,58	9	0	tidak
9,78	9	55,87	9	0	tidak
9,55	8	56,71	9	0	tidak
6,38	9,61	68,5	9	0	tidak
6,61	9,67	67,64	9	0	tidak
7,44	9,89	64,23	9	0	tidak
5,11	9,61	73,66	9	0	tidak
8,28	9,56	61,04	9	0	tidak
6,67	9,61	67,32	9	0	tidak
8,77	10	59,4	9	0	tidak
8,16	9,78	61,5	9	0	tidak
6,17	9,61	69,3	9	0	tidak
7,33	8,94	64,73	9	0	tidak
7,67	9,5	63,34	9	0	tidak
7,22	8,78	65,07	9	0	tidak
8,5	8,61	60,29	9	0	tidak
6,95	9,39	66,21	10	0	tidak
6	8,5	70,06	10	0	tidak
6,5	10,6	67,99	10	0	tidak
6,61	9,61	67,53	10	8,89	ringan
7,05	9,44	65,8	10	199,9	sangatlebat
9,83	8,72	55,71	10	0	tidak
7,34	9,72	64,66	10	0	tidak
3,94	5,11	79,04	10	0	tidak
4,05	4	78,59	10	17,02	ringan
6,56	8	67,68	10	3,05	ringan
8,12	7,78	61,62	10	0	tidak
4,72	5,89	75,41	10	0	tidak
5,28	7,39	73,24	10	0	tidak
6,45	8,28	68,37	10	0	tidak
6,39	9	68,61	10	0	tidak
6,5	8	68,13	10	0	tidak
5,23	8	73,28	10	0	tidak
3,11	6,78	83,2	10	16	ringan
5,5	7,78	72,39	10	0	tidak
7,34	7,78	64,99	10	2,03	ringan

B-8

TD	RT	RH	B L N	PRC P	Class	TD	RT	RH	B L N	PRC P	Class
7,94	8,28	62,6	10	0	tidak	3,33	6,22	82,06	12	0	tidak
8	9,5	62,38	10	0	tidak	2,33	7,11	87,06	12	20,83	sedang
7,83	9,72	62,82	10	0	tidak	2,55	9	85,88	12	47,24	sedang
5,38	7,89	72,6	10	0	tidak	4,83	9,61	75,12	12	1,02	ringan
4,89	8,11	74,83	11	0	tidak	2,78	8	84,75	12	27,94	sedang
6,17	8,67	69,48	11	0	tidak	4,5	7,28	76,6	12	0	tidak
5,44	8,89	72,54	11	0	tidak	4,83	8,39	75,09	12	0	tidak
7,28	8,5	65,12	11	0	tidak	5,17	7,72	73,67	12	0	tidak
6,27	8,78	69,15	11	0	tidak	4,5	7,72	76,63	12	2,03	ringan
7,11	9,22	65,72	11	0	tidak	4,66	7,22	75,88	12	0	tidak
4,94	8,33	74,69	11	0	tidak	4,06	7,5	78,57	12	0	tidak
6,67	8	67,61	11	0	tidak	3,28	7,11	82,31	12	7,11	ringan
6,39	7,28	68,68	11	0	tidak	4,22	9,61	77,85	12	1,52	ringan
7,34	8	64,89	11	0	tidak	3,56	8,78	80,92	1	0	tidak
6,06	7,83	69,93	11	0	tidak	4,55	9,78	76,32	1	91,19	lebat
8,38	7,89	61,06	11	0	tidak	4,61	9,22	76,14	1	7,87	ringan
6,34	7,83	68,79	11	0	tidak	4,39	7,89	77,06	1	0	tidak
3,72	8,5	80,26	11	0,76	ringan	3,78	6,89	79,97	1	0	tidak
2,72	6,78	85,08	11	6,1	ringan	5,11	8,39	73,88	1	13,97	ringan
3,73	8,39	80,1	11	9,91	ringan	4,39	8,11	77,04	1	0	tidak
3	7,5	83,73	11	1,02	ringan	6	7,78	70,22	1	0	tidak
2,94	6,89	83,99	11	4,06	ringan	4,06	8,22	78,64	1	0	tidak
4,61	6,83	76,21	11	0	tidak	3,55	7,78	81,01	1	4,06	ringan
5,44	6,61	72,66	11	0	tidak	3,73	8,11	80,23	1	0	tidak
4,78	8,22	75,44	11	0	tidak	3,05	8	83,46	1	0,76	ringan
5,06	8,78	74,21	11	0,76	ringan	2,95	5,78	84,01	1	95	lebat
3,94	10	79,14	11	0,51	ringan	3,5	6,39	81,29	1	1,02	ringan
4,33	7,5	77,34	11	0	tidak	3	7	83,63	1	1,02	ringan
4,84	7,67	75,07	11	0	tidak	3,84	8,89	79,55	1	35,05	sedang
5,89	7,94	70,63	11	35,05	sedang	3,66	11,8	80,46	1	23,88	sedang
3,72	8,11	80,14	11	0,76	ringan	1,66	7,89	90,6	1	1,02	ringan
4,06	7,28	78,48	12	0	tidak	1,55	6,11	91,14	1	23,11	sedang
5,17	8,22	73,54	12	0	tidak	3,77	7,39	80,01	1	1,02	ringan
5,72	9	71,26	12	0	tidak	2,67	7	85,3	1	23,11	sedang
7,17	9,33	65,35	12	0	tidak	2,89	8,78	84,18	1	30,99	sedang
5,89	9	70,45	12	0,76	ringan	2,05	5,56	88,39	1	2,03	ringan
6,11	9	69,43	12	0	tidak	3,89	12	79,41	2	3,05	ringan
5,33	8,39	72,61	12	0	tidak	3,66	6,39	80,45	2	0	tidak
6,33	8	68,57	12	0	tidak	3,16	8	82,87	2	16	ringan
5,44	8,39	72,37	12	0	tidak	3,45	7,17	81,45	2	0,76	ringan

TD	RT	RH	B L N	PRC P	Class
3,61	8,11	80,63	2	0,76	ringan
5,72	8,61	71,11	2	0	tidak
2	5	88,71	2	0,51	ringan
4,61	6,17	76,12	2	1,02	ringan
5,44	9,39	72,4	2	3,05	ringan
3,5	8	81,13	2	27,18	sedang
3,56	5	80,85	2	0,76	ringan
4,17	9,61	78,01	2	0,76	ringan
4,34	7	77,29	2	0	tidak
4,44	9,61	76,83	2	0	tidak
5,78	9	71	2	0	tidak
3,44	9,11	81,48	2	5,08	ringan
3,61	6,89	80,59	2	8,89	ringan
4,11	8,44	78,39	2	0	tidak
3,95	8	79,17	2	0,76	ringan
3,83	8,11	79,72	2	0	tidak
4,22	7,22	77,96	3	0	tidak
3,16	7,78	82,86	3	0	tidak
6,06	7,61	69,96	3	0	tidak
4,11	7,61	78,39	3	0	tidak
3,33	7,89	82,05	3	0	tidak
3,05	8,22	83,41	3	0,76	ringan
4,05	8,28	78,73	3	0	tidak
6	8,39	70,23	3	0	tidak
4,84	7	75,15	3	0	tidak
3,73	7,22	80,23	3	0	tidak
3,77	6,39	80,06	3	0,51	ringan
3,5	7,39	81,29	3	0	tidak
7,33	7,61	64,85	3	0	tidak
4,72	8	75,77	3	0	tidak
5,45	8,22	72,63	3	0	tidak
5,89	8	70,74	3	0	tidak
3,78	7,61	80,01	3	0	tidak
4,89	8,61	75,05	3	0	tidak
2,89	8	84,21	3	0	tidak
3,33	8,56	82,11	3	0,51	ringan
3,11	6,94	83,21	3	0,51	ringan
2,56	6,72	85,92	3	11,94	ringan
2,89	4,61	84,29	3	4,06	ringan
5,83	6	70,83	4	0	tidak

TD	RT	RH	B L N	PRC P	Class
3,78	8,39	79,86	4	0,51	ringan
4,83	7,61	75,28	4	0	tidak
7,23	8	65,28	4	0,51	ringan
4,45	9	76,75	4	0	tidak
4,33	7,39	77,5	4	0	tidak
3,17	8	82,88	4	0	tidak
4,5	6,11	76,67	4	11,94	ringan
4,88	7,11	74,99	4	0	tidak
4,84	7,44	75,13	4	0	tidak
3,94	6,67	79,21	4	0	tidak
3,78	7	79,95	4	0	tidak
3,22	7,61	82,58	4	0,51	ringan
4,06	7,17	78,68	4	0	tidak
5,95	6,78	70,44	4	0	tidak
5	7,5	74,48	4	4,06	ringan
6,11	7,89	69,76	4	0	tidak
5,61	8	71,83	4	0	tidak
6,5	7	68,1	4	0	tidak
6,5	7,61	68,18	4	0	tidak
7,33	7,5	64,86	4	0	tidak
6,45	8,22	68,28	4	0	tidak
6,39	8	68,59	5	0	tidak
5,95	7,72	70,52	5	0	tidak
5,17	8,17	73,76	5	0,76	ringan
5,84	6,72	71	5	0	tidak
5,45	6,89	72,63	5	0	tidak
4,72	7,39	75,72	5	0	tidak
5,28	7,33	73,33	5	0	tidak
5,33	7	73,06	5	0	tidak
3,06	7,72	83,38	5	9,91	ringan
4,55	7	76,57	5	0	tidak
4,22	6,89	78,01	5	0	tidak
5,5	9,11	72,43	01	0	tidak
4,39	6,06	77,27	01	0	tidak
4,5	7,83	76,7	01	10,9	ringan
4	7,72	79,12	01	0	tidak
4,44	5,83	77,04	01	0	tidak
5,78	6,89	71,26	01	0	tidak
6,61	8,33	67,92	01	0	tidak
5,44	8,5	72,68	01	0	tidak

B-10

TD	RT	RH	B L N	PRC P	Class
5,62	8,33	71,92	01	1,02	ringan
4,72	7,22	75,76	01	3,05	ringan
4,5	6,11	76,77	01	0	tidak
5,39	8,39	72,89	01	0	tidak
5,84	6,39	71,07	01	7,87	ringan
5,83	8,11	71,18	01	0	tidak
5,89	8,61	70,87	01	1,02	ringan
6,44	7,5	68,52	01	0	tidak
4,33	8,22	77,51	01	32	sedang
3,11	7,5	83,28	01	27,9	sedang
3,62	7,78	80,81	01	0	tidak
2,17	6,39	87,95	01	19,1	ringan
3,05	7,5	83,49	01	13	ringan
3,28	7,28	82,42	01	1,02	ringan
2,22	5,28	87,65	01	17,5	ringan
2,39	4,61	86,75	01	3,05	ringan
2,34	7,39	86,99	01	6,35	ringan
2,61	6,5	85,62	01	0	tidak
3,23	9,39	82,49	01	0,25	ringan
6,06	10,5	70	01	0	tidak
5,33	10,3 9	73,14	01	0,76	ringan
4,28	8,39	77,83	01	4,06	ringan
3,45	7,78	81,6	01	2,03	ringan
4,28	7,5	77,72	02	0	tidak
3,06	7,28	83,45	02	11,9	ringan
4,94	6,17	74,65	02	0	tidak
2,11	7,5	88,22	02	15	ringan
3,56	8,06	81,1	02	0	tidak
2,5	6	86,23	02	0,51	ringan
2,06	6,72	88,51	02	83,1	lebat
2,95	7,06	84,01	02	25,9	sedang
2,11	6,28	88,24	02	55,9	lebat
2,33	5,5	87,08	02	2,79	ringan
2,72	4,61	85,07	02	1,02	ringan
4,11	7,33	78,53	02	0,25	ringan
4,78	7,44	75,52	02	5,08	ringan
2,33	4,11	87,14	02	12,2	ringan
2,89	5,22	84,3	02	43,9	sedang
4,05	9,22	78,7	02	18	ringan

TD	RT	RH	B L N	PRC P	Class
4,17	8,61	78,3	02	3,05	ringan
4,72	7,72	75,84	02	0	tidak
3,67	6	80,56	02	16	ringan
2,22	6,61	87,68	02	31	sedang
3,67	8,61	80,51	02	0	tidak
2,61	7,5	85,64	02	35,1	sedang
2,5	8,5	86,18	02	51,3	lebat
3	9,11	83,75	02	5,08	ringan
2,61	7,5	85,64	02	32	sedang
2,61	7,78	85,63	02	24,9	sedang
3,78	8,22	79,99	02	0	tidak
4,17	7,61	78,19	02	0	tidak
3,22	7,5	82,67	03	6,1	ringan
2,61	8	85,68	03	39,9	sedang
3,39	7,61	81,9	03	0	tidak
2,39	11,1 1	86,86	03	4,06	ringan
2,83	7,44	84,58	03	3,05	ringan
4,16	8,28	78,34	03	0	tidak
4,11	7,17	78,59	03	0	tidak
3,94	7	79,4	03	0	tidak
3,94	8,22	79,45	03	13	ringan
3,28	8,28	82,44	03	9,91	ringan
4,28	9,78	77,75	03	0	tidak
3,95	7,61	79,32	03	0,25	ringan
3,67	7,78	80,58	03	11,9	ringan
5,17	7,83	73,79	03	0	tidak
3,84	8,61	79,82	03	15	ringan
4,33	6,78	77,57	03	0	tidak
3	6,56	83,81	03	26,9	sedang
3,61	6,89	80,93	03	0	tidak
4,56	7,28	76,63	03	0	tidak
3,89	7,39	79,6	03	24,9	sedang
4,05	7,78	78,84	03	0	tidak
4,61	8,11	76,31	03	0,51	ringan
3,66	8,67	80,67	03	23,1	sedang
4,56	8,44	76,52	03	0	tidak
4,67	8	76,02	03	1,02	ringan
3,45	6,5	81,63	03	0	tidak
3,33	6,72	82,21	03	7,87	ringan

TD	RT	RH	B L N	PRC P	Class
2,727		85,19	03	20,6	sedang
3,338,39		82,15	03	0	tidak
2,615,33		85,71	04	56,9	lebat
4,226,39		78,1	04	0	tidak
4,388,22		77,35	04	6,1	ringan
3,679,06		80,64	04	18	ringan
3,777,61		80,21	04	0	tidak
3,285		82,53	04	0,25	ringan
3,177,89		83,04	04	46	sedang
2,566,56		85,98	04	22,1	sedang
3	6,5	83,82	04	1,02	ringan
3,676,5		80,68	04	5,08	ringan
3	6	83,83	04	0,76	ringan
3,116,22		83,25	04	0,76	ringan
3,727,39		80,39	04	6,1	ringan
2,053,78		88,61	04	14,5	ringan
4,396,39		77,31	04	0	tidak
3,235,28		82,71	04	0	tidak
3,786,56		80,12	04	0	tidak
4,126,39		78,5	04	1,02	ringan
4,228,5		78,01	04	3,05	ringan
4,557,33		76,52	04	2,03	ringan
3,345,61		82,14	04	0	tidak
5,116,72		74,12	04	0	tidak
4,735,61		75,85	04	0	tidak
5,726		71,57	04	0	tidak
4,786,72		75,47	04	0	tidak
5,617,72		71,91	04	0	tidak
4,946,61		74,89	04	0	tidak
4,776,89		75,63	04	0	tidak
4,726		75,88	04	0,25	ringan
3,895,78		79,62	05	14	ringan
3,945,89		79,4	05	0	tidak
3,057		83,63	05	46	sedang
4,226,5		78,11	05	0	tidak
4,567,44		76,62	05	0	tidak
3,895,72		79,62	05	0,51	ringan
3,287		82,49	05	41,7	sedang
3,119,28		83,21	05	0,76	ringan
3	6	83,88	05	22,1	sedang

TD	RT	RH	B L N	PRC P	Class
4,116,61		78,61	05	0	tidak
3,286,72		82,5	05	3,05	ringan
4,566,72		76,62	05	0	tidak
5,116,83		74,26	05	0	tidak
4,726,61		75,96	05	1,02	ringan
4,5	7,22	76,88	05	0	tidak
4,067,78		78,82	05	69,1	lebat
4,787,83		75,55	05	0	tidak
5,837,22		71,03	05	0	tidak
4,566,72		76,62	05	0	tidak
3,5	6,61	81,43	05	17	ringan
3,786,78		80,09	05	0	tidak
4,836,11		75,37	05	0	tidak
4,956,72		74,8	05	11,9	ringan
5,555,22		72,23	05	0	tidak
5,565,83		72,24	05	0	tidak
3,896,11		79,46	05	0	tidak
3,666		80,67	05	1,02	ringan
3,565,39		81,16	05	10,9	ringan
2,177,22		87,91	05	79	lebat
3,618,22		80,74	05	0	tidak
3,5	6	81,39	06	3,05	ringan
5,346,39		73,05	06	0	tidak
5,395,78		72,78	06	0	tidak
5,457,72		72,54	06	0	tidak
4,616,39		76,27	06	0	tidak
4,286,61		77,76	06	5,08	ringan
3,897,22		79,62	06	0,76	ringan
5,286,61		73,32	06	0	tidak
4,456,06		76,99	06	0	tidak
5,116,39		74,08	06	0	tidak
5,676		71,66	06	0	tidak
5,336		73,13	06	0	tidak
4,226,17		78,03	06	0	tidak
3,5	5,5	81,34	06	0	tidak
5,057,11		74,31	06	0	tidak
3,176,72		82,99	06	23,9	sedang
2,897,83		84,27	06	23,9	sedang
2,617,61		85,58	06	0	tidak
4,5	8,11	76,63	06	0	tidak

B-12

TD	RT	RH	B L N	PRC P	Class	TD	RT	RH	B L N	PRC P	Class
3,95	6,72	79,26	06	0	tidak	4,11	7	78,46	08	2,03	ringan
4,06	7	78,62	06	21,1	sedang	4,11	5,89	78,54	08	0	tidak
3,67	7,72	80,52	06	5,08	ringan	5,11	5,67	73,93	08	0	tidak
4,72	7,39	75,7	06	0	tidak	5	8	74,3	08	0	tidak
5,05	6,17	74,23	06	0	tidak	5,89	8,72	70,47	08	0	tidak
5	7,11	74,42	06	0,25	ringan	5,28	8,11	73,22	08	0,25	ringan
3,89	7,17	79,47	06	2,03	ringan	4,06	5,61	78,64	08	1,02	ringan
3,67	5,89	80,46	06	7,11	ringan	5,73	6,39	71,2	08	0	tidak
2,22	5,39	87,72	06	35,1	sedang	5,39	6,5	72,66	08	0	tidak
1,72	4	90,27	06	0,25	ringan	5,27	6,78	73,26	08	0	tidak
3,56	7	81	07	0	tidak	4,95	6,5	74,64	08	0	tidak
4	6,61	79,04	07	0	tidak	5,5	7,72	72,12	08	0	tidak
4,34	6,56	77,48	07	0	tidak	7,06	7,39	65,51	08	0	tidak
5,56	6,11	72,1	07	0	tidak	6,94	8,5	65,96	08	0	tidak
4,95	5,28	74,7	07	0	tidak	7,23	8,72	64,75	08	0	tidak
4,83	5,61	75,25	07	0	tidak	7	8,28	65,63	08	0	tidak
5	6,61	74,41	07	0	tidak	6,28	8,72	68,6	08	0	tidak
6,34	8	68,73	07	0	tidak	5,89	7,72	70,3	08	0	tidak
5,05	6,22	74,28	07	0	tidak	6,22	7,61	69,01	08	0	tidak
3,67	5,83	80,58	07	0	tidak	5,28	6,5	73,17	08	0	tidak
3,72	6,11	80,34	07	0	tidak	4,34	6,67	77,49	08	6,1	ringan
3,61	6,5	80,87	07	21,1	sedang	5,05	6,28	74,28	08	0	tidak
3,5	5,22	81,36	07	41,9	sedang	5,16	5,11	73,83	08	0	tidak
3	4	83,75	07	4,83	ringan	5,39	9,83	72,83	08	0	tidak
2,89	6	84,33	07	0,51	ringan	6,17	6,39	69,5	08	0	tidak
2,61	6	85,65	07	8,89	ringan	6	8	70,2	09	0	tidak
5,06	7,5	74,13	07	0,51	ringan	6,5	8,17	68,09	09	0,51	ringan
1,89	5,89	89,37	07	8,38	ringan	5,17	8	73,65	09	1,02	ringan
3,72	5,78	80,29	07	3,05	ringan	5,61	7,67	71,8	09	0	tidak
4,83	6,39	75,07	07	0	tidak	5,66	5,5	71,64	09	0	tidak
5,33	7,28	72,85	07	0	tidak	5,67	6,11	71,64	09	0	tidak
5,56	7,39	71,87	07	0,76	ringan	6,22	6,83	69,25	09	0	tidak
5,72	7	71,26	07	0	tidak	7,17	8	65,41	09	0	tidak
5,67	6,5	71,52	07	0	tidak	6,94	7,28	66,34	09	0	tidak
6,17	5,72	69,35	07	0	tidak	6,72	8,5	67,19	09	0	tidak
4,56	7	76,25	07	0	tidak	7,61	8	63,59	09	0	tidak
5,39	6,61	72,53	08	0	tidak	6,33	7,78	68,62	09	0	tidak
6,5	7,61	67,7	08	0	tidak	6,16	8,39	69,38	09	0	tidak
5,61	7,39	71,6	08	0	tidak	6,67	8,61	67,35	09	0	tidak
5	7,11	74,42	08	0	tidak	5,11	7	74,05	09	0	tidak

TD	RT	RH	B L N	PRC	Class
4,94	6,11	74,75	09	0	tidak
5	5,39	74,57	09	1,02	ringan
4,94	6,11	74,82	09	0	tidak
5,22	7,22	73,57	09	0	tidak
4,16	9,78	78,23	09	0	tidak
2,56	9,17	85,89	09	4,06	ringan
3,73	11,78	80,22	09	0,25	ringan

Halaman ini sengaja dikosongkan

LAMPIRAN C

HASIL UJI PERFORMA CART

C.1 Akurasi Training (CART) -Pengendalian kedalaman pohon keputusan

Tabel C-1 Nilai akurasi Training (CART) -Pengendalian kedalaman pohon keputusan

	Min Parent							
Min leaf	1	3	5	10	30	50	70	90
1	1	0,9848	0,9658	0,9321	0,8687	0,8504	0,838	0,8289
3	0,9308	0,9308	0,9308	0,9098	0,8631	0,8466	0,8357	0,8286
5	0,8958	0,8958	0,8958	0,8958	0,8598	0,8401	0,8312	0,8256
10	0,8624	0,8624	0,8624	0,8624	0,8504	0,8365	0,8276	0,8248
20	0,8352	0,8352	0,8352	0,8352	0,8352	0,8324	0,8271	0,8243
30	0,8274	0,8274	0,8274	0,8274	0,8274	0,8274	0,8264	0,8236
40	0,8243	0,8243	0,8243	0,8243	0,8243	0,8243	0,8243	0,8218
50	0,8218	0,8218	0,8218	0,8218	0,8218	0,8218	0,8218	0,8218
60	0,819	0,819	0,819	0,819	0,819	0,819	0,819	0,819
70	0,818	0,818	0,818	0,818	0,818	0,818	0,818	0,818
80	0,8139	0,8139	0,8139	0,8139	0,8139	0,8139	0,8139	0,8139
90	0,8139	0,8139	0,8139	0,8139	0,8139	0,8139	0,8139	0,8139
100	0,8132	0,8132	0,8132	0,8132	0,8132	0,8132	0,8132	0,8132

C.2 Akurasi Testing (CART) -Pengendalian kedalaman pohon keputusan

Tabel C-2 Nilai akurasi testing (CART) -Pengendalian kedalaman pohon keputusan

	Min Parent							
Min leaf	1	3	5	10	30	50	70	90
1	0,7126	0,7132	0,7161	0,7286	0,7516	0,7617	0,7635	0,77

90	0,304	0,304	0,304	0,304	0,304	0,304	0,304	0,304
100	0,298	0,298	0,298	0,298	0,298	0,298	0,298	0,298

C.4 Nilai Presisi kelas 0 (CART)

Tabel C-4 Nilai Presisi Kelas 0 (CART)

	Min Parent							
Min leaf	1	3	5	10	30	50	70	90
1	0,863	0,867	0,867	0,870	0,869	0,859	0,870	0,866
3	0,867	0,867	0,867	0,871	0,868	0,858	0,869	0,866
5	0,861	0,861	0,861	0,861	0,866	0,864	0,871	0,868
10	0,869	0,869	0,869	0,869	0,868	0,866	0,874	0,868
20	0,871	0,871	0,871	0,871	0,871	0,877	0,874	0,868
30	0,879	0,879	0,879	0,879	0,879	0,879	0,871	0,866
40	0,870	0,870	0,870	0,870	0,870	0,870	0,870	0,859
50	0,859	0,859	0,859	0,859	0,859	0,859	0,859	0,859
60	0,859	0,859	0,859	0,859	0,859	0,859	0,859	0,859
70	0,854	0,854	0,854	0,854	0,854	0,854	0,854	0,854
80	0,857	0,857	0,857	0,857	0,857	0,857	0,857	0,857
90	0,857	0,857	0,857	0,857	0,857	0,857	0,857	0,857
100	0,878	0,878	0,878	0,878	0,878	0,878	0,878	0,878

C.5 Nilai Presisi kelas 1 (CART)

Tabel C-5 Nilai Presisi kelas 1 (CART)

	Min Parent							
Min leaf	1	3	5	10	30	50	70	90
1	0,356	0,356	0,364	0,385	0,417	0,445	0,437	0,447
3	0,361	0,361	0,361	0,414	0,422	0,438	0,437	0,445
5	0,410	0,410	0,410	0,410	0,416	0,442	0,443	0,451
10	0,426	0,426	0,426	0,426	0,438	0,450	0,452	0,455

20	0,436	0,436	0,436	0,436	0,436	0,452	0,459	0,464
30	0,473	0,473	0,473	0,473	0,473	0,473	0,465	0,470
40	0,462	0,462	0,462	0,462	0,462	0,462	0,462	0,470
50	0,470	0,470	0,470	0,470	0,470	0,470	0,470	0,470
60	0,478	0,478	0,478	0,478	0,478	0,478	0,478	0,478
70	0,473	0,473	0,473	0,473	0,473	0,473	0,473	0,473
80	0,454	0,454	0,454	0,454	0,454	0,454	0,454	0,454
90	0,454	0,454	0,454	0,454	0,454	0,454	0,454	0,454
100	0,447	0,447	0,447	0,447	0,447	0,447	0,447	0,447

C.6 Nilai Presisi kelas 2 (CART)

Tabel C-6 Nilai Presisi kelas 2 (CART)

	Min Parent							
Min leaf	1	3	5	10	30	50	70	90
1	0,356	0,356	0,364	0,385	0,417	0,445	0,437	0,447
3	0,361	0,361	0,361	0,414	0,422	0,438	0,437	0,445
5	0,410	0,410	0,410	0,410	0,416	0,442	0,443	0,451
10	0,426	0,426	0,426	0,426	0,438	0,450	0,452	0,455
20	0,436	0,436	0,436	0,436	0,436	0,452	0,459	0,464
30	0,473	0,473	0,473	0,473	0,473	0,473	0,465	0,470
40	0,462	0,462	0,462	0,462	0,462	0,462	0,462	0,470
50	0,470	0,470	0,470	0,470	0,470	0,470	0,470	0,470
60	0,478	0,478	0,478	0,478	0,478	0,478	0,478	0,478
70	0,473	0,473	0,473	0,473	0,473	0,473	0,473	0,473
80	0,454	0,454	0,454	0,454	0,454	0,454	0,454	0,454
90	0,454	0,454	0,454	0,454	0,454	0,454	0,454	0,454
100	0,447	0,447	0,447	0,447	0,447	0,447	0,447	0,447

C.7 Nilai Presisi kelas 4 (CART)

Tabel C-7 Nilai Presisi kelas 4 (CART)

	Min Parent
--	------------

[illegible]

C.8 Nilai Recall kelas 0 (CART)

Tabel C-8 Nilai Recall kelas 0 (CART)

[illegible]

80	0,915	0,915	0,915	0,915	0,915	0,915	0,915	0,915
90	0,915	0,915	0,915	0,915	0,915	0,915	0,915	0,915
100	0,904	0,904	0,904	0,904	0,904	0,904	0,904	0,904

C.9 Nilai Recall kelas 1 (CART)

Tabel C-9 Nilai Recall kelas 1 (CART)

	Min Parent							
Min leaf	1	3	5	10	30	50	70	90
1	0,416	0,416	0,413	0,445	0,448	0,416	0,513	0,487
3	0,403	0,403	0,403	0,448	0,429	0,406	0,506	0,487
5	0,439	0,439	0,439	0,439	0,426	0,429	0,516	0,494
10	0,448	0,448	0,448	0,448	0,465	0,439	0,526	0,494
20	0,487	0,487	0,487	0,487	0,487	0,500	0,548	0,516
30	0,571	0,571	0,571	0,571	0,571	0,571	0,539	0,506
40	0,526	0,526	0,526	0,526	0,526	0,526	0,526	0,481
50	0,481	0,481	0,481	0,481	0,481	0,481	0,481	0,481
60	0,458	0,458	0,458	0,458	0,458	0,458	0,458	0,458
70	0,445	0,445	0,445	0,445	0,445	0,445	0,445	0,445
80	0,465	0,465	0,465	0,465	0,465	0,465	0,465	0,465
90	0,465	0,465	0,465	0,465	0,465	0,465	0,465	0,465
100	0,584	0,584	0,584	0,584	0,584	0,584	0,584	0,584

C.10 Nilai Recall kelas 2 (CART)

Tabel C-10 Nilai Recall kelas 2 (CART)

	Min Parent							
Min leaf	1	3	5	10	30	50	70	90
1	0,119	0,119	0,119	0,129	0,139	0,198	0,119	0,109
3	0,158	0,158	0,158	0,188	0,178	0,198	0,119	0,109
5	0,208	0,208	0,208	0,208	0,178	0,188	0,109	0,109
10	0,139	0,139	0,139	0,139	0,139	0,188	0,109	0,109
20	0,168	0,168	0,168	0,168	0,168	0,198	0,099	0,099

[illegible]

C.11 Nilai Recall kelas 4 (CART)

Tabel C-11 Nilai Recall kelas 4 (CART)

[illegible]

C-8

Halaman ini sengaja dikosongkan

LAMPIRAN D

HASIL UJI PERFORMA MODEL RANDOM FOREST

D.1 Akurasi Training (Random Forest) - Pengendalian parameter

Tabel D-1 Akurasi Training (Random Forest) - Pengendalian parameter

	Jumlah Pohon											
Variabel Acak	10	20	30	40	50	80	100	150	200	250		
1	0,989	0,997	0,998	1	0,999	1	1	1	1	1		
2	0,99	0,998	0,999	0,999	1	1	1	1	1	1		
3	0,991	0,999	0,999	0,999	1	1	1	1	1	1		
4	0,993	0,998	0,999	1	1	1	1	1	1	1		
5	0,991	0,998	1	1	1	1	1	1	1	1		
6	0,991	0,998	1	0,999	1	1	1	1	1	1		
7	0,991	0,999	0,998	1	0,999	1	1	1	1	1		
8	0,989	0,997	0,999	0,999	1	1	1	1	1	1		
9	0,989	0,998	0,999	0,999	1	1	1	1	1	1		
10	0,992	0,998	1	0,999	1	1	1	1	1	1		
11	0,99	0,994	1	1	1	1	1	1	1	1		
12	0,991	0,995	0,998	0,999	1	1	1	1	1	1		

D.2 Akurasi Testing (Random Forest) - Pengendalian parameter

Tabel D-2 Akurasi Testing (Random Forest) - Pengendalian parameter

	Jumlah Pohon											
Var Acak	10	20	30	40	50	80	100	150	200	250		

1	0,766	0,781	0,783	0,782	0,785	0,785	0,784	0,780	0,788	0,782
2	0,769	0,774	0,785	0,782	0,788	0,788	0,788	0,785	0,781	0,789
3	0,775	0,781	0,778	0,778	0,779	0,779	0,786	0,783	0,782	0,785
4	0,769	0,780	0,776	0,779	0,777	0,777	0,785	0,784	0,784	0,785
5	0,765	0,785	0,782	0,780	0,781	0,781	0,786	0,783	0,779	0,784
6	0,777	0,781	0,774	0,787	0,773	0,773	0,788	0,782	0,784	0,782
7	0,759	0,778	0,777	0,776	0,777	0,777	0,783	0,781	0,784	0,782
8	0,771	0,774	0,781	0,781	0,778	0,778	0,785	0,784	0,784	0,785
9	0,763	0,770	0,774	0,782	0,779	0,779	0,785	0,778	0,778	0,777
10	0,756	0,777	0,781	0,777	0,783	0,783	0,780	0,782	0,779	0,777
11	0,764	0,777	0,774	0,778	0,778	0,778	0,779	0,784	0,782	0,781
12	0,769	0,775	0,775	0,775	0,779	0,779	0,778	0,784	0,781	0,784

D.3 Weighted Akurasi Testing (Random Forest) - Pengendalian parameter

Tabel D-3Nilai Weighted Akurasi Testing (Random Forest) -
Pengendalian parameter

	Jumlah Pohon									
Var Acak	10	20	30	40	50	80	100	150	200	250
1	0,303	0,302	0,307	0,308	0,295	0,31	0,302	0,309	0,303	0,305
2	0,322	0,3	0,317	0,316	0,305	0,31	0,306	0,309	0,313	0,318
3	0,309	0,312	0,305	0,314	0,322	0,303	0,31	0,316	0,323	0,312
4	0,318	0,309	0,323	0,31	0,308	0,307	0,309	0,305	0,311	0,315
5	0,314	0,321	0,318	0,318	0,309	0,316	0,31	0,317	0,304	0,312
6	0,345	0,32	0,307	0,332	0,321	0,312	0,314	0,313	0,305	0,313
7	0,315	0,309	0,313	0,315	0,326	0,308	0,311	0,315	0,317	0,317
8	0,314	0,321	0,32	0,316	0,309	0,311	0,311	0,314	0,316	0,314
9	0,323	0,299	0,314	0,321	0,315	0,314	0,309	0,321	0,311	0,308
10	0,317	0,308	0,319	0,31	0,299	0,32	0,316	0,306	0,314	0,316

11	0,31	0,319	0,301	0,311	0,322	0,313	0,312	0,309	0,307	0,309
12	0,333	0,31	0,309	0,303	0,318	0,312	0,319	0,316	0,324	0,315

D.4 Nilai Presisi kelas 0 (Random Forest)

Tabel D-4 Nilai Presisi kelas 0 (random forest)

	Jumlah Pohon										
Var Acak	10	20	30	40	50	80	100	150	200	250	
1	0,867	0,874	0,879	0,871	0,869	0,875	0,871	0,876	0,878	0,872	
2	0,875	0,873	0,873	0,875	0,876	0,879	0,876	0,876	0,877	0,881	
3	0,872	0,875	0,877	0,874	0,874	0,875	0,880	0,877	0,878	0,878	
4	0,876	0,876	0,879	0,872	0,873	0,874	0,875	0,872	0,878	0,878	
5	0,873	0,879	0,880	0,879	0,882	0,872	0,880	0,878	0,881	0,882	
6	0,875	0,882	0,878	0,875	0,882	0,874	0,878	0,880	0,881	0,883	
7	0,878	0,882	0,873	0,878	0,883	0,874	0,879	0,874	0,879	0,880	
8	0,879	0,879	0,878	0,874	0,879	0,877	0,877	0,875	0,879	0,879	
9	0,871	0,876	0,877	0,877	0,877	0,872	0,879	0,881	0,883	0,880	
10	0,875	0,877	0,877	0,875	0,877	0,882	0,880	0,879	0,878	0,879	
11	0,880	0,871	0,874	0,877	0,882	0,882	0,879	0,878	0,877	0,877	
12	0,873	0,873	0,873	0,876	0,877	0,884	0,876	0,876	0,877	0,878	

D.5 Nilai Presisi kelas 1 (Random Forest)

Tabel D-5 Nilai Presisi kelas 1 (Random Forest)

	Jumlah Pohon										
Var Acak	10	20	30	40	50	80	100	150	200	250	
1	0,442	0,474	0,474	0,470	0,442	0,472	0,463	0,474	0,462	0,466	
2	0,454	0,462	0,478	0,469	0,464	0,477	0,481	0,477	0,469	0,482	
3	0,481	0,472	0,467	0,450	0,457	0,473	0,481	0,486	0,464	0,475	
4	0,452	0,478	0,446	0,467	0,460	0,458	0,479	0,460	0,473	0,468	
5	0,438	0,481	0,469	0,461	0,457	0,455	0,480	0,479	0,475	0,472	
6	0,471	0,458	0,441	0,488	0,480	0,457	0,476	0,473	0,471	0,471	

D-4

7	0,426	0,468	0,468	0,455	0,476	0,469	0,463	0,452	0,468	0,476
8	0,460	0,451	0,476	0,467	0,449	0,474	0,480	0,458	0,458	0,470
9	0,436	0,446	0,460	0,483	0,471	0,451	0,468	0,478	0,472	0,467
10	0,421	0,460	0,477	0,463	0,446	0,480	0,470	0,460	0,468	0,464
11	0,447	0,471	0,453	0,467	0,474	0,469	0,480	0,459	0,467	0,457
12	0,452	0,460	0,454	0,458	0,467	0,471	0,460	0,460	0,479	0,473

D.6 Nilai Presisi kelas 2 (Random Forest)

Tabel D-6 Nilai Presisi kelas 2 (Random Forest)

	Jumlah Pohon									
Var Acak	10	20	30	40	50	80	100	150	200	250
1	0,333	0,259	0,280	0,313	0,313	0,529	0,400	0,429	0,429	0,400
2	0,324	0,212	0,370	0,375	0,304	0,355	0,333	0,417	0,400	0,409
3	0,286	0,364	0,244	0,379	0,379	0,448	0,414	0,448	0,320	0,400
4	0,326	0,294	0,333	0,355	0,343	0,394	0,393	0,333	0,370	0,345
5	0,319	0,372	0,389	0,367	0,333	0,290	0,371	0,364	0,375	0,367
6	0,392	0,382	0,269	0,395	0,313	0,323	0,361	0,333	0,353	0,361
7	0,267	0,275	0,273	0,306	0,382	0,333	0,355	0,313	0,323	0,361
8	0,294	0,282	0,367	0,379	0,343	0,371	0,429	0,316	0,367	0,344
9	0,349	0,220	0,341	0,324	0,333	0,286	0,382	0,333	0,313	0,333
10	0,316	0,324	0,344	0,333	0,233	0,273	0,371	0,353	0,342	0,333
11	0,264	0,361	0,231	0,308	0,353	0,243	0,353	0,306	0,324	0,273
12	0,320	0,270	0,333	0,257	0,316	0,310	0,324	0,306	0,314	0,342

D.7 Nilai Presisi kelas 4 (Random Forest)

Tabel D-7 Nilai Presisi kelas 4 (Random Forest)

	Jumlah Pohon									
Var Acak	10	20	30	40	50	80	100	150	200	250

1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

D.8 Nilai Recall kelas 0 (Random Forest)

Tabel D-8 Nilai Recall kelas 0 (Random Forest)

	Jumlah Pohon									
Var Acak	10	20	30	40	50	80	100	150	200	250
1	0,905	0,923	0,918	0,921	0,918	0,921	0,918	0,924	0,919	0,925
2	0,900	0,914	0,922	0,916	0,918	0,922	0,925	0,926	0,922	0,920
3	0,908	0,918	0,921	0,917	0,915	0,919	0,927	0,923	0,918	0,919
4	0,902	0,916	0,911	0,917	0,921	0,920	0,927	0,918	0,921	0,919
5	0,902	0,918	0,914	0,915	0,913	0,918	0,921	0,920	0,918	0,921
6	0,906	0,916	0,912	0,922	0,918	0,919	0,920	0,918	0,921	0,922
7	0,895	0,915	0,919	0,912	0,918	0,917	0,917	0,918	0,915	0,922
8	0,904	0,908	0,914	0,919	0,914	0,915	0,919	0,916	0,917	0,920
9	0,894	0,913	0,906	0,915	0,915	0,912	0,914	0,922	0,917	0,918
10	0,887	0,916	0,914	0,911	0,918	0,918	0,919	0,918	0,918	0,916
11	0,895	0,912	0,917	0,914	0,919	0,913	0,915	0,914	0,918	0,920
12	0,901	0,918	0,916	0,915	0,917	0,909	0,914	0,918	0,922	0,918

D.9 Nilai Recall kelas 1 (Random Forest)**Tabel D-9 Nilai Recall kelas 1 (Random Forest)**

	Jumlah Pohon									
Var Acak	10	20	30	40	50	80	100	150	200	250
1	0,494	0,519	0,548	0,535	0,506	0,548	0,539	0,548	0,552	0,529
2	0,523	0,516	0,529	0,542	0,535	0,535	0,532	0,532	0,532	0,568
3	0,539	0,526	0,503	0,510	0,516	0,532	0,535	0,542	0,539	0,548
4	0,516	0,532	0,516	0,523	0,500	0,506	0,519	0,519	0,539	0,535
5	0,487	0,529	0,539	0,529	0,545	0,506	0,535	0,539	0,545	0,542
6	0,516	0,523	0,519	0,523	0,552	0,510	0,532	0,539	0,532	0,532
7	0,490	0,523	0,494	0,519	0,548	0,510	0,532	0,503	0,542	0,532
8	0,516	0,519	0,542	0,519	0,513	0,535	0,535	0,510	0,529	0,532
9	0,503	0,494	0,526	0,539	0,529	0,503	0,539	0,535	0,552	0,532
10	0,487	0,513	0,535	0,529	0,510	0,552	0,529	0,519	0,523	0,526
11	0,516	0,519	0,497	0,519	0,539	0,545	0,545	0,526	0,526	0,510
12	0,503	0,497	0,490	0,513	0,519	0,548	0,519	0,513	0,526	0,529

D.10 Nilai Recall kelas 2 (Random Forest)**Tabel D-10 Nilai Recall kelas 2 (Random Forest)**

	Jumlah Pohon									
Var Acak	10	20	30	40	50	80	100	150	200	250
1	0,119	0,069	0,069	0,050	0,050	0,089	0,059	0,059	0,059	0,529
2	0,119	0,069	0,099	0,089	0,069	0,109	0,099	0,099	0,099	0,568
3	0,099	0,119	0,099	0,109	0,109	0,129	0,119	0,129	0,079	0,548
4	0,139	0,099	0,119	0,109	0,119	0,129	0,109	0,089	0,099	0,535
5	0,149	0,158	0,139	0,109	0,089	0,089	0,129	0,119	0,119	0,542
6	0,198	0,129	0,069	0,149	0,099	0,099	0,129	0,109	0,119	0,532

